VISIT…

LANZAROTE
Caliente.com

Series of Linguistics
复旦博学 · 语言学系列

博學



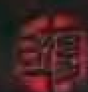
Essentials of Applied Linguistics

# 应用语言学纲要

齐沪扬　陈昌来　主编

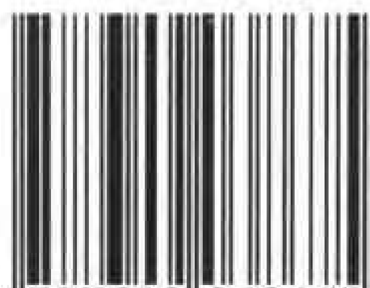
复旦大学出版社
www.fudanpress.com.cn



复旦博学 · 语言学系列

责任编辑 韩结根
封面设计 马晓霞


ISBN 7-309-03867-3
9 787309 038675 >
H · 770 定价: 28.00元

复旦








Essentials of Applied Linguistics

# 应用语言学纲要

主　编　齐沪扬　陈昌来

编写者　肖奚强　张亚军　姚占龙　胡建锋　张小峰

复旦大學出版社

图书在版编目(CIP)数据

应用语言学纲要/齐沪扬、陈昌来主编.—上海:复旦大学出版社,2004.4
ISBN 7-309-03867-3

Ⅰ.应… Ⅱ.①齐…②陈… Ⅲ.应用语言学 Ⅳ.H08

中国版本图书馆CIP数据核字(2004)第001222号

应用语言学纲要
齐沪扬 陈昌来 主编

出版发行 复旦大学出版社
上海市国权路579号 邮编200433
86-21-65118853(发行部) 86-21-65109143(邮购)
fupnet@fudanpress.com http://www.fudanpress.com

责任编辑 韩结根
装帧设计 马晓霞
总 编 辑 高若海
出 品 人 贺圣遂

印　　刷 上海长阳印刷厂
开　　本 787×960 1/16
印　　张 20.25
字　　数 331千
版　　次 2004年4月第一版 2004年4月第一次印刷
印　　数 1—6 000

书　　号 ISBN 7-309-03867-3/H·770
定　　价 28.00元

如有印装质量问题,请向复旦大学出版社发行部调换。
版权所有 侵权必究

## 内容提要

本书为应用语言学概论性教材。全书吸取当前应用语言学研究的最新成果，系统阐述了应用语言学的产生、发展、性质、学科地位与研究方法，并集中介绍了语言教学、对外汉语教学、语言测试、中文信息处理、语言规划和语言调查、社会语言学、儿童语言学、地名学和人名学等应用语言学的八个主要研究领域，体现了应用语言学的实用性和针对性。

本书既可作为对外汉语系、中文系、语言学系、应用语言学系的“应用语言学”课程的教材，也可供语言学、应用语言学及汉语言文字学等专业的研究生作为基础课教材，还可供外语系本科生和研究生以及其他对应用语言学有兴趣的学人参考。

# 前　言

语言是人类最重要的交际工具，语言伴随着人类的产生而产生，也随着人类社会的发展而发展，正因为如此，人类对语言的研究也很早就开始了。在语言研究中，人们很早就关注语言应用的有关问题，如语言教学（包括第二语言教学）的理论和实践，字母及文字的创制和选择，正字法的确立和规范，标准语的确立和规范，字典和词典的编纂，语言与社会文化关系的探讨等。研究语言应用的种种问题的学问就是应用语言学（Applied linguistics）。虽然对语言应用问题的研究起源较早，但作为一门学科，“应用语言学”的形成却是很晚的。跟应用语言学相对的是理论语言学（也称一般语言学或普通语言学）和专门语言学（或叫本体语言学，如现代汉语、古代汉语、现代英语，或现代汉语语法学、现代汉语词汇学、现代英语语音学等）。早期的语言研究中，这三者的研究是融合在一起的，如研究某种具体的语言，需要语言理论，也要关注具体语言的应用。随着历史比较语言学，尤其是结构主义语言学的兴起和发展，旨在研究不同语言的一般问题的理论语言学（一般语言学或普通语言学）迅速发展①。19 世纪以后的语言学理论研究和应用研究开始分化，如语言教学跟当时的语言历史问题的研究（历史比较语言学）就已经分家。19 世纪后期，1870 年波兰语言学家博杜恩·德·库尔特内提出要区分“纯粹语言学”和“应用语言学”，首次提出“应用语言学”这个术语，但并没有引起人们的注意。直到 20 世纪 40 年代，由于外语教学的发展，人们才开始重视应用语言学的研究。一般认为美国是应用语言学的发源地。二战前后，由于军事、科技、文教、旅游以及政治、经济的需要，美国开始大力发展外语教学（尤其是对外英语教学）。1946 年美国密执安大学建立了英语学院，研究如何对外国人教授英语，并出版著名的杂志《语言学习》

① 有人认为德国语言学家洪堡特（1767—1835 年）是普通语言学的奠基人。

(Language Learning),这个刊物的副标题就是《应用语言学杂志》(Journal of Applied Linguistics),这是世界上第一本应用语言学杂志,而且主要是研究语言教学问题的。1959 年美国华盛顿成立"应用语言学中心",主要开展英语和外语教学研究。另外,1958 年英国爱丁堡大学研究生部开始建立应用语言学学院,以培养应用语言学的专门人才,随后,世界上不少大学都建立了应用语言学专业。标志应用语言学学科正式形成的是 1964 年在法国召开的第一届国际应用语言学大会和国际应用语言学协会的成立。国际应用语言学会议到目前为止已经召开 13 届(一般每 3 年一次)。应用语言学目前在国外已经成为语言学的一门重要分支学科。

我国对语言应用问题的研究起源于先秦,历代的学者在语言教学理论和实践,语言文字规划、词典编纂等方面都取得了一定的成就。但中国的应用语言学学科的形成却更晚,1984 年成立了语言文字应用研究所(成立时属于中国社会科学院和中国文字改革委员会共同领导,后归国家语言文字工作委员会,1998 年归属于教育部),这标志着应用语言学在中国正式形成。1992 年专门的应用语言学杂志《语言文字应用》创刊,1995 年筹建了中国应用语言学学会(筹),该学会主持召开数次全国性的应用语言学学术研讨会,同时出版了数种应用语言学论著。近几年来,许多大学或学术机构开始招收语言学及应用语言学的研究生(包括硕士研究生和博士研究生),同时,作为应用语言学的重要分支学科如语言规划、对外汉语教学、外语教学、计算语言学等近年来也取得了突出的发展等等,这些都标志着中国的应用语言学的进一步发展和趋于成熟。

对应用语言学的理解有广义和狭义两种,狭义的应用语言学专指语言教学,特指外语教学和第二语言教学,这种理解的应用语言学可以定义为研究语言理论在语言教学或外语教学中的种种应用问题,这也是应用语言学发展最早和最为充分的一个分支。广义的应用语言学是语言学在各方面的应用以及各有关实际领域的学问的统称,即指语言学知识和研究成果所应用的一切领域和方面,应用语言学所关心的是如何应用语言学理论、方法和成果来阐释其他应用领域所遇到的跟语言有关的问题。由此可见,应用语言学是语言学跟其他学科相互交叉渗透所产生的一门边缘学科①。

应用语言学作为一门学科从 20 世纪 60 年代发展到今天,已经是一门

① 随着应用语言学的发展和语言学应用领域的拓展,目前多采取广义的理解。

比较成熟的语言学分支学科，人们对应用语言学的认识也越来越明确。总起来看，应用语言学有如下一些特点。

首先，应用语言学具有学科的相对独立性。应用语言学是语言学中一门相对独立的学科，表现在它有明确的研究任务，研究语言学在一切领域的实际应用问题，有明确的研究对象，有像语言教学、语言规划、语言信息处理、神经语言学、社会语言学、心理语言学、词典学等几个较为成熟的下位领域，有大量的研究论著、教材、专门的研究刊物、国际性学术组织、众多的专门研究人才，有专门的应用语言学专业和课程，以培养高学位的后备人才。

其次，应用语言学具有实用性的特点。实用性是应用语言学存在和发展的基本条件，应用语言学就是要解决语言学在各个应用领域中的实际问题，如指导和帮助语言教学、语言故障康复、语言信息处理、语言规划、词典编纂、翻译等等。应用语言学是为了直接满足语言学在社会生活中的实际需求。

跟实用性相关的是应用语言学具有实验性的特点。应用语言学要解决语言运用的实际问题就离不开调查和实验，如进行语言教学或第二语言教学要对教学对象、语言自身的特点、中介语现象、教学效果等方面进行调查和分析，对语言教学的新方法是否有效要进行相关的实验，提取必要的分析数据，才能得出科学的结论；进行语言规划要对语言文字自身的特点和现状、语言文字使用者的状况、语言规划的目的和效果等问题进行调查和研究；进行社会语言学研究更要进行必要的社会调查；进行语言信息处理研究必须懂得使用计算机做各种实验。调查法和实验法是应用语言学研究的两个重要的方法。调查法是应用语言学研究的常用方法，包括访谈调查、观察调查、问卷调查（包括网上问卷调查）等。实验法也是神经语言学、语言教学、计算语言学、社会语言学等领域的常用研究方法。无论是调查法还是实验法，都要对材料、数据、结论进行统计、比较和分析，因而统计的手段在应用语言学中较为常用，利用统计的方法来进行研究有利于使研究的结论达到定量和定性的统一，以保证结论的可靠性和科学性。

强调应用语言学的实用性和实验性，绝不是要否定应用语言学的理论特点，任何学科都必须建立在一定的理论基础之上，关于应用语言学的性质、特点、任务、范围、对象、研究方法等基本理论和各具体领域的理论和方法的研究，是应用语言学不可缺少的部分，没有理论和方法的指导，应用语言学是不可能成为一门新兴学科的，没有理论和方法以及研究手段的更新和发展，应用语言学也不能取得进一步的发展。不过，相对理论语言学和普

通语言学来说，应用语言学的实用性和实验性的特点更加突出一些。

应用语言学还具有综合性的特点，这是由应用语言学的学科性质决定的。应用语言学要根据具体的研究对象和研究目的同其他学科结合，研究应用语言学不仅需要语言学知识，也需要相关学科的知识，如语言教学研究要吸取教育学、心理学、教育测量学、学科教学论等学科的理论和方法；研究语言规划离不开政治学、民族学的指导；研究社会语言学需要社会学、文化学、人类学等学科的理论和方法；研究计算语言学当然要跟计算机科学、数理逻辑、人工智能、信息论等学科结合。正因为应用语言学在不同领域跟不同学科的结合，才产生了应用语言学的许多下位学科，如社会语言学、心理语言学、神经语言学、病理语言学、计算语言学、人类语言学、语言风格学等等，因而学习和研究应用语言学，除了需要语言学的知识和理论方法外，还需要更多其他学科的知识，应用语言学的研究人员应该是一种复合型的人才，培养应用语言学人才也需要多学科的合作。

应用语言学的研究对象既包括语言教学①、语言文字规划、词典学、人名学和地名学等传统的研究领域，也包括社会语言学、心理语言学、神经语言学、病理语言学、计算语言学、人类语言学、语言风格学、语言测试、术语学、计算语言学(包括语言信息处理)、儿童语言习得、翻译学、言语交际学、传播语言学(广播、电影、电视、报刊等新闻媒体语言以及广告语言、网络语言等)、法律语言学、刑侦语言学、体态语等研究领域。总之，对应用语言学的研究对象应该持一种开放的态度，随着语言应用领域的扩展，应用语言学的研究领域也必将随之拓展。

本书将介绍应用语言学的八个主要研究领域，它们分别是：语言教学、对外汉语教学、语言测试、中文信息处理、语言规划和语言调查、社会语言学、儿童语言学、地名学和人名学。选择上述这些研究领域主要是基于以下考虑：一是这些领域在应用语言学学科中发展相对成熟，有比较好的研究基础，取得了较多的成果，它们也是应用语言学学科的主要研究对象。二是考虑到集中介绍几个主要问题，有利于在有限的篇幅内把问题说明得更加系

① 作为应用语言学的最传统分支的语言教学，内涵已经十分广泛，目前，在我国语言教学包括母语教学、外语教学即第二语言教学、对外汉语教学、少数民族语言教学、双语教育、特殊人群如聋哑人和弱智人群的语言教学、计算机辅助语言教学等许多部门，其中外语教学、对外汉语教学已经成为我国应用语言学的重要领域。

统全面一些，应用语言学的领域是开放的，众多的领域不可能在一本书中都介绍到，否则，虽然面面俱到，却只能蜻蜓点水，点到为止，对学习者来说反而不得要领。最后，选择这些领域更主要的是考虑到本书的适用对象。本书主要是供中文系、对外汉语系、语言学系、应用语言学系本科生使用，可以作为这些本科专业的教材；同时本书也可以供语言学及应用语言学、汉语言文字学等专业的研究生作为基础课的教材。随着社会发展的需要，社会上急需掌握应用语言学基本知识的人才，因而在对外汉语专业、汉语言专业、语言学专业、汉语言文学专业的本科教育中，急需开设应用语言学课程，需要应用语言学的教材；随着应用语言学的发展，急需培养应用语言学研究的高级人才，因而语言学及应用语言学、汉语言文字学专业的研究生也需要学习应用语言学的课程。就这些专业的学生来说，本书所介绍的八个研究领域，相对来说更具有针对性和实用性。当然，本书也可以供外语系本科生和研究生以及其他对应用语言学有兴趣的学人参考。

## 参考书目

1. 冯志伟. 应用语言学综论. 广东教育出版社，1999
2. 龚千炎、冯志伟. 应用语言学研究刍议，语文建设 1991(6)
3. 桂诗春. 应用语言学. 湖南教育出版社，1988
4. 刘涌泉、乔毅. 应用语言学. 上海教育出版社，1988
5. 〔英〕S·皮特·科德. 应用语言学导论. 上海外语教育出版社，1988
6. 许嘉璐. 语言文字学及其应用研究. 广东教育出版社，1999
7. 于根元. 二十世纪的中国语言应用研究. 书海出版社，1996
8. 于根元主编. 应用语言学纲要. 华语教学出版社，1999
9. 于根元主编. 世纪之交的应用语言学. 北京广播学院出版社，2000
10. 于根元主编. 应用语言学概论. 商务印书馆，2003

# 目　　录


前言 …………………………………………………… 1

第一章　语言教学 ……………………………………… 1

第一节　语言教学的性质和基本过程 …………………… 1

一、语言教学的性质和任务 ……………………………… 1

二、语言教学的目标和类型 ……………………………… 2

三、语言教学的基本过程 ………………………………… 4

第二节　汉语文教学的发展历史和思考 ………………… 7

一、萌芽时期 ……………………………………………… 7

二、传统时期 ……………………………………………… 8

三、变革时期 ……………………………………………… 12

四、改革实验时期 ………………………………………… 14

五、语文教学的繁荣发展和思考 ………………………… 15

第三节　国外语言教学的发展 …………………………… 18

一、古代和中世纪 ………………………………………… 18

二、文艺复兴时期 ………………………………………… 19

三、17 世纪和 18 世纪 …………………………………… 20

四、19 世纪 ……………………………………………… 20

五、20 世纪 ……………………………………………… 22

六、教学方法的构成和类型 ……………………………… 25

第四节　语言理论 ………………………………………… 26

一、概念的妥当性 ………………………………………… 26

二、语言的诸方面 ………………………………………… 29

三、术语 …………………………………………………… 32

**第二章 对外汉语教学** …… 34
第一节 学科概况 …… 34
一、学科性质与名称 …… 34
二、学科的历史与现状 …… 38
第二节 教学大纲、课程设置及教材编写 …… 43
一、教学大纲 …… 43
二、课程设置 …… 44
三、教材编写 …… 46
第三节 教学流派和课堂教学 …… 49
一、第二语言教学流派简介 …… 49
二、课堂教学 …… 56
第四节 第二语言习得 …… 61
一、对比分析理论 …… 61
二、偏误分析和中介语理论 …… 65

**第三章 语言测试** …… 69
第一节 实验方法在语言学中的应用 …… 69
一、定量方法和定性方法 …… 69
二、实验方法的意义和特点 …… 71
三、测量的信度和效度等问题 …… 74
第二节 现代语言测试的理论框架 …… 77
一、语言测试在语言教学中的地位 …… 77
二、语言测试的性质 …… 80
三、语言测试的两大根本要求——信度和效度 …… 83
第三节 语言测试中试题的生产程序 …… 85
一、试题的设计和制作 …… 85
二、试题的评分 …… 87
三、试题的质量控制和统计分析 …… 89
第四节 语言测试中不同题型的具体命题问题 …… 92
一、试题和题型 …… 92
二、MC 类题的命题 …… 92
三、阅读理解题的命题 …… 95

四、写作题的命题 …… 98

**第四章　中文信息处理** …… 105
第一节　中文信息处理概述 …… 105
一、中文信息处理及其研究内容 …… 105
二、国内中文信息处理的发展 …… 110
三、中文信息处理面临的挑战 …… 113
第二节　中文自动分词和自动词性标注 …… 116
一、中文自动分词概述 …… 116
二、歧义字段与自动分词 …… 120
三、自动词性标注 …… 124
第三节　中文句处理 …… 129
一、中文句处理概述 …… 129
二、歧义短语及其排歧策略 …… 133
三、句处理中的语义问题 …… 138

**第五章　语言规划与语言调查** …… 150
第一节　语言规划 …… 150
一、语言规划的定义、性质和原则 …… 150
二、语言文字的地位规划和本体规划 …… 153
三、文字改革 …… 159
四、汉语拼音运动与汉语拼音方案 …… 160
五、普通话的推广 …… 161
六、汉字的整理、简化和标准化 …… 166
七、术语的标准化 …… 169
第二节　语言调查 …… 171
一、基本程序和方法 …… 171
二、资料收集 …… 176
三、资料整理 …… 179

**第六章　社会语言学** …… 182
第一节　社会语言学的性质 …… 182

一、语言与社会交际 …… 182
二、社会语言学 …… 185
三、社会语言学研究和流派 …… 189
第二节 宏观社会语言学 …… 192
一、语言影响 …… 192
二、语言转用 …… 195
三、语言兼用——双语现象 …… 197
四、语言混合——皮钦语和克里奥尔语 …… 201
五、语言冲突和语言地位规划 …… 203
第三节 微观社会语言学 …… 206
一、语言变异和语言变体 …… 206
二、语言的地域变体 …… 207
三、语言的社会变体 …… 212

**第七章 儿童语言发展** …… 226
第一节 儿童语言发展的阶段与儿童语言获得理论 …… 226
一、儿童语言发展研究概述 …… 226
二、儿童语言发展的阶段和规律 …… 229
三、儿童语言获得理论 …… 235
第二节 儿童语音的发展 …… 241
一、儿童声音、语音的听辨 …… 242
二、儿童语音发生的准备 …… 243
三、儿童音位系统的发展 …… 245
第三节 儿童词汇的发展 …… 247
一、不同年龄段儿童的词汇量 …… 248
二、儿童词汇习得过程中各类词的发展状况 …… 249
三、儿童词义的习得 …… 252
四、儿童词汇、语义发展的特点 …… 254
第四节 儿童语法的发展 …… 255
一、句子的平均长度 …… 255
二、不完整句阶段 …… 257
三、完整句阶段 …… 260

**第八章　地名学和人名学** …… 268
第一节　概述 …… 268
一、地名和地名学 …… 268
二、人名和人名学 …… 273
第二节　汉语地名学 …… 278
一、地名词语的语言学特征 …… 278
二、地名词语系统 …… 283
三、地名规范化 …… 287
第三节　汉语人名学 …… 291
一、汉语人名的语言学特征 …… 291
二、汉语人名系统 …… 297
三、人名的规范 …… 306

**后记** …… 309

# 第一章 语言教学

语言教学跟应用语言学的关系最为密切，狭义的应用语言学似乎专指语言教学，尤其是指第二语言教学或外语教学[1]。如英国学者科德(S. Pit Corder)的《应用语言学导论》就是专门介绍语言教学的[2]。语言教学涉及的内容和问题较为广泛，它跟语言学、语言习得、心理学、教育学、学科教学论、教育技术、教育测量等学科都有密切的关系。本章主要介绍语言教学的一般问题和汉语文教学、国外语言教学的主要问题。对外汉语教学、第二语言教学和语言测试等目前的热门话题，将在第二章和第三章中单独介绍。

## 第一节 语言教学的性质和基本过程

### 一、语言教学的性质和任务

语言教学(language teaching)是指一种有目的、有计划、有特定方法的教学活动，是教育工作的重要组成部分。语言教学作为教学活动就是为了培养和提高受教育者听、说、读、写的语言能力。虽然一个人从一岁前后就开始牙牙学语，他的父母或周围的其他人不断地教他说话，进入幼儿园以后，保教人员也开始通过游戏、图片、故事等一定的手段来训练他的听和说，实际上一个孩子在上小学之前已经能用母语比较熟练地表达和交流。但孩子无论在家学习语言或在幼儿园学习语言，总的来说，其语言学习还是属于自然习得，语言学习过程是不自觉的或无计划的，因而严格来说，幼儿的语言学习或对幼儿的语言训练还不是

---

① 《语言学百科词典》，上海辞书出版社，1993

② S·皮特·科德《应用语言学导论》，上海外语教育出版社，1983

正式的语言教学活动①。

儿童入小学以后，其语言(语文)教学才是有计划、有组织、有目标的教学活动，在一定的教学大纲的指导下，来系统地培养和提高学生的语言能力——听、说、读、写能力，尤其是提高学生的书面表达能力，并在此基础上培养和提高学生的文学、文化水平和思维能力、道德情操。

可见，语言教学是指关于语言的有目的的、有计划的、有特定教学方法的教学活动。

语言教学有悠久的历史。在中国，3 000 多年前就有了语言教学，先秦时期政府就设置了学校，《大戴礼记·保傅篇》说："及太子少长……则入于小学，小学者所学之宫也。……古者八岁而就外舍，学小艺焉。""小艺"有 6 种，其中之一是"六书"，就是关于语言文字的学习，当然也就有了语言文字的教学。当时不仅有正规的童蒙学习语言文字的课本，如《史籀篇》及后代的《仓颉篇》等，而且有了一定的教学理念，如《礼记·学记》、《荀子·劝学篇》等论著就对教与学的方法和技巧给予了一定的总结。先秦时的孔子授徒就有了"文学"科和"语言"科的分别。在西方，语言教学也起源很早，古希腊和古罗马时代，为了教授古代经典，就开始了语言教学，如亚历山大时代，希腊语教育得到了全面促进，公元前 3 世纪前后马其顿的统治者在诸如小亚细亚的珀加蒙和埃及的亚历山大里亚等城市里，建立并资助使用希腊语的大学，那里较高级的教学和研究的内容包括希腊文学、荷马时代的古代希腊语言等。欧洲中世纪时，学校教育分"三才"、"四技"等不同学科，所谓"三才"即三学科，指语法、逻辑(即雄辩术)、修辞。可见，语言教学(教学拉丁语)在欧洲中世纪的学校教育中占有重要的地位。

语言教学除了指教学活动外，还指研究语言教学活动的全过程和各个具体环节的理论、原则和方法的科学，即语言教学研究。

## 二、语言教学的目标和类型

### 1. 语言教学的目标

语言教学的总目标是培养和提高受教育者的语言能力。"语言能力"传

① 目前有的幼儿园甚至托儿所已经开始正规的语言训练，有的幼儿园就有语言课，但主要以口语为主。

统上指运用语言进行听、说、读、写的能力，或指采用口语或书面语进行交际的能力。这种意义上的“语言能力”主要包括有关母语或第二语言的语音、语义(词汇)、语法以及语句篇章衔接的知识，对于书面语交际来说，还包括记录母语的文字方面的知识。具备这些知识可以保证受教育者规范地、正确地使用语言。但这还远远不够，一个人真正的语言能力还应包括恰当地、得体地、有效地、成功地使用语言的能力。后一种意义上的语言能力一般称为语言交际能力，主要指说话者适切交际主体、主旨、语境有效地进行言语交际的能力。可见，语言教学的目标不仅要教授受教育者正确地、规范地使用语言的能力，还要教授受教育者恰当地、得体地、有效地、成功地运用语言的能力。传统的语言教学多是教授语言的组成规则，培养学生规范使用语言的能力，而现代语言教学则更加强调语言交际能力的培养。现代语言教学不仅要教授语言的组成规则，还要教授说话的规则，这样，语言教学中就要考虑到语言运用的诸多因素，如语境、社会文化背景知识等，要考虑到言语交际的各种规则，如合作原则、礼貌原则等。

2. 语言教学的类型

就语言教学的具体情况来看，有的语言教学是教授受教育者学习和运用本族语的系统知识进行口语和书面语交际，同时学习和领会本族语所负载的本民族的风俗习惯、文化传统、道德情操。这就是第一语言教学或母语教学(first language teaching, mother tongue teaching)。第一语言教学也就是传统的语文教学，它一般开始于正规的学校教育，通常是从小学阶段正式开始。语文教学应该是听、说、读、写同步进行，但书面语教学如阅读教学和写作教学在传统的语文教学中占有更重要的地位，因为小学生已经能熟练运用母语进行口语交际。不过，语文教学中，不能忽视口语训练，口语教学应该是语文教学不可缺少的部分。考虑到当前社会交际领域的日益扩大，社会交际的方式、内容日益复杂，口语交际的日益重要，今后的语文教学中要适当的、有针对性的进行口语交际能力的教学和训练，如演讲训练、应聘求职的语言技巧训练、上下级之间或同事之间的语言沟通训练、如何与陌生人有效沟通的语言技巧训练、社会调查和采访的语言训练、日常生活交往的语言能力训练、电话交谈或网上交谈的语言文明教育、自身口头语言素质的提高等等，这些都应该是语文教学中语文素质教育的应有内容。第一语言教学不仅要传授语言知识，使学生获得驾驭母语进行听、说、读、写的各种交际和表达能力，而且要通过语文学习获得丰富的文化知识、生活常识，并

受到道德的教育、情操的熏陶。

有的语言教学是在受教育者已经或基本掌握母语的基础上再教授受教育者掌握第二语言的知识和运用第二语言的能力。这就是非本族语教学，也叫第二语言教学(second language teaching)。第二语言教学的目标是培养学生运用第二语言进行口语或书面语的交际能力。不过，第二语言的教学目标往往因人而异，有的学习者学习第二语言是为了看外语文献，教学中就应重视阅读训练；有的学习者是为了日常交际，教学中就应该重视口语训练；有的仅仅是为了旅游，有的仅仅是为了做生意，则只需要注重旅游语言、旅游文化或经贸语言；有的是为了融入第二语言所在的社会，则需要进行全面的语言知识和语言能力的训练。

第二语言教学包括外语教学和少数民族语言教学。对外汉语教学属于第二语言教学中的外语教学。少数民族语言教学在我国主要指对少数民族的汉语教育，基本上属于双语教育的范畴，也包括在少数民族地区工作的汉族干部或其他人员学习少数民族语言。

### 三、语言教学的基本过程

语言教学是一个复杂的过程，它涉及到语言学、心理学、教育学和哲学等众多社会科学，也涉及教学技术、计算机科学等自然科学，更跟政治、经济以及一个国家的教育政策有密切的关系，所以，语言教学不仅仅指课堂教学活动，它是一个多层次的综合过程，这一过程要由政府、应用语言学家、任课教师及学生来共同完成。

一般来说，语言教学包括如下几个阶段：

#### 1. 语言教学政策的制订

语言教学首先要由政府或政府主管部门(如教育部)制定政策性的语言规划，解决要不要教语言，教哪种语言(多语国家或多民族国家尤其要首先解决这个问题，另外，选择第二语言教学也需要政府指导)，教谁语言等问题；政府及政府各部门还需要确定语言教学的目的，提供必要的教育经费和教育设施，帮助培训教师和管理人员。这一阶段是属于政策性的。

#### 2. 语言教学的总体设计

其次是由语言学家或应用语言学家或其他专业人员来具体执行上述政策，主要是进行语言教学的总体设计，如确定教学要求、教学内容和教学时间，确定课程设置及各门课程的具体教学方式，制定教学大纲和教学计划，

编写或选择教材。语言教学的总体设计在语言教学的全过程中有着至关重要的作用,它不仅要解决教学目标和教学要求,而且要解决教什么(教学内容)、什么时候教(教学对象)、教多少(教学顺序)、用什么(教材)教、如何教(教学原则和教学方法)等问题,还包括课时安排、教学阶段划分、课程设置等具体问题。总体设计是否科学合理,直接关系到语言教学的成败。总体设计是对整个语言教学过程的宏观把握,它能针对不同教学对象来帮助教学者找到符合语言学习规律和语言教学规律、符合教学要求和客观条件的最佳教学方案,协调好各个教学环节,使整个教学过程和全部教学活动成为一个统一的、科学的整体。因而,语言教学,无论是第一语言教学,还是第二语言教学,都应该重视语言教学的总体设计。

3. 课堂教学

第三个阶段是由任课教师实际施教,即课堂教学。课堂教学是语言教学的关键,语言教学政策和总体设计的具体实施者是任课教师,语言教学的成功与否在于任课教师的具体的课堂教学活动。任课教师必须具备丰富的语言知识和语言理论,有一定的科学文化修养,了解语言教学的目标,掌握科学的语言教学方法并善于运用,运用好现代教育技术,具有娴熟和灵活的课堂教学技巧,并且要善于组织教学。课堂教学不仅指课堂教学活动,还应该包括有组织、有计划的课外语言实践。

4. 语言测试

最后一个阶段是语言测试,因为上述所有环节都要在学生身上实施,并要通过测试和错误分析来检查语言教学的具体效果。因而,语言教学的评价与测试是语言教学过程的重要环节,通过对教学过程中诸要素的评价和对学生的测试,来检验语言教学活动的成效,以便对语言教学的各项计划及其执行情况作出评价和反馈。目前的测试方法主要有主观测试法和客观测试法,两种测试各有利弊,最好是将主观和客观结合起来。而测试的具体形式主要有3种:一是“鉴别性测试”或“分类测试”,主要用于入学或分班测试;二是“成绩测试”或“学业测试”,主要用于检查学生课堂学习成绩;三是“能力测试”或“水平测试”,用于测试学生或受测试者对某种语言的掌握情况和实际水平。语言测试的项目应该包括听、说、读(包括语法、词汇、理解)、写、译等多个方面,同时不同学习目的要有不同侧重。

语言测试是语言教学的重要一环。目前的语言测试已经进入科学化、标准化、实用化、多样化、国际化阶段,许多国家都对语言测试进行了大量的

研究，一些语言如英语的测试已经在国际上推出了许多种类型，如托福考试、GRE考试、雅思考试、剑桥英语系列考试等，中国的英语考试也有中小学英语等级考试、大学英语四级六级考试、专业英语四级八级考试等，各国英语考试还推出口语等级、口译等级、商务英语等类型考试。汉语水平考试（HSK）是检测母语非汉语学习者汉语水平的最权威考试类型，目前已经有相当高的国际影响；而且有关部门和机构正在研制和开发专业性汉语水平考试，如旅游HSK、商贸HSK、少儿HSK、文秘HSK等。语言测试的形式也从主观性考试到主观和客观相结合考试，从传统考试方式到现代化、标准化考试，当前一些语言测试已经开始利用计算机技术采用自适应性考试类型。在测试研究上，应用语言学家和教育测量学家已经做了大量工作，对考试的信度、效度、区分度、难易度以及实用性、可靠性、针对性进行了许多研究。

5. 语言教学的两个核心问题：语言教学大纲和语言教学法

语言教学是一个复杂的动态的系统，但在这个复杂过程中，"教什么"和"如何教"是最为基本的问题，也是语言教学的核心。"教什么"就是制定科学合理完善的教学大纲。语言教学大纲是语言教学过程的全面计划，它根据不同的教学目的和要求，对教学内容进行组织和安排，以便在编写教材时并在进行教学活动时有所依据。可见，语言教学大纲是语言教学活动的核心环节，主要是要解决教什么的问题。不过，大纲仅仅是提出语言教学的目标，要达到所规定的目标，还要研究和制定语言教学的具体步骤，确定实施各项步骤的具体途径，因而大纲跟教材编写和教学方法是紧密联系的，即"教什么（大纲）—用什么教（教材）—如何教（教学法）"三者缺一不可。过去在中学语法教学的争论中，由于过分重视对语法体系（即语法大纲）的论证，而忽视了编写出适合的中学语法教材，更忽视了对如何教好中学语法的研究，以致我国的中学语法教学一直没有取得良好的收效。

语言教学大纲应该在语言学、社会语言学、教育学、教学法、心理学等众多学科的指导下进行编写，但主要要以语言学原则为基础。语言教学大纲的编写首先要确定项目的选择，即根据一定的培养目标来选择语言的语音、词汇、语法、修辞、文字等要素项目。其次是根据教学对象和教学方法对项目的组织和编排，例如是以形式为标准编排项目，还是以意义或功能为标准来编排项目，哪些项目在前，哪些项目在后，等等。大纲编排的标准往往跟教学法相联系。如20世纪初的语言大纲是以单词为纲的编排法，50年代出现了以结构和句型为纲的编排法，60年代产生了结构编排法，70年代以

后出现了功能-意念大纲。功能-意念大纲不仅包括了传统的词汇和结构，而且还包含功能、意念、场景、社会及心理因素等等，它既考虑到语言交际能力，也兼顾了语法和语境因素。

“如何教”即教学法是语言教学的又一个核心环节，语言教学法既指具体的教学方式和方法，更指整个语言教学的方法体系。后者在第二语言教学即外语教学中尤其受重视，近100多年来形成了不同的语言教学法体系，按出现的先后顺序大体有语法翻译法、直接法、听说法、视听法、功能法等。我国在第一语言教学即语文教学中也总结出许多有用的方法，如识字教学中总结出“分散识字”、“集中识字”、“注音识字，提前读写”等教学法，尤其“集中识字，大量阅读，分步写作”识字教学法已经把识字、阅读、写作融为一体，不仅仅是单纯的识字教学了。

## 第二节 汉语文教学的发展历史和思考

我国的语文教学有着悠久的历史，有文字记载的汉语文教育的历史至少有3 000多年了，我们将汉语文教育的发展过程分为萌芽时期、传统时期、变革时期、改革实验时期、繁荣发展的新时期5个阶段。

### 一、萌芽时期

我国的语文教学萌芽于殷周到春秋战国时期。在那个时代教育已经相当发达，先后建立了官学和私学。官学中有“小学”和“大学”，许慎《说文解字》“叙”中说“周礼八岁入小学，保氏教国子先以六书”，据段玉裁注引《大戴礼记·保傅篇》说“古者年八岁，而就外舍，学小艺焉，履小节焉。束发而就大学，学大艺焉，履大节焉”。虽然，段玉裁注中所引文献对入小学和大学的年龄说明有所不同，但至少可以说明“小学”是为当时的贵族子弟(国子，即公卿大夫之子弟)所设立的初级学校，“大学”则是高级学校。孔子首开私学之风气，设馆授徒。有了学校就有了教育理念，先秦时期已经有了比较成熟和比较完备的教育和教学思想，如《礼记》中的《学记》和荀子的《劝学》以及《论语》中就包含了古人的丰富的教学理念，如《学记》中提出“相观而善”的互相切磋的方法，《劝学》中提出“善假于物”的学习方法，荀子《效儒》提出积

累学习的途径：故圣人者，人之所积也。《论语》中更是提出了“博学”、“慎思”、“温故知新”、“启发”式、“问一知十、举一反三”、“由博返约综合贯通”、“有教无类”、“因材施教”等等教与学的方法。先秦的教育内容应该是以语言文字和文学的教学为主要内容，如以“六书”为识字教学之始，孔子授徒就有“文学”和“言语”两科。先秦是非常重视语文能力的培养，提倡善于言辞，孔子就曾说过不学《诗》就可能不善于言辞这样的话。从屈原、宋玉等人的文学作品、诸子百家的著作、几部史书(《春秋》、《左传》、《战国策》、《国语》等)所记载的各诸侯国互相派遣的使节的辞令来判断，当时的语文教育是相当发达并且是卓有成效的。

### 二、传统时期

秦汉至清末是中国语文教育的传统时期，在这近 2 000 年的历史发展中，虽然不同时期都会有所变化，但总的来说都有较为一致的共性，在语文教学中逐渐形成了一定的范式。传统时期的语文教学从教学内容来看主要可以分为识字教学、书法教学、读书教学、作文教学等几个方面。

#### 1. 识字教学

传统语文教学主要是教授学生学习古代经典，即学习文言文和写作文言文。由于时代的变迁和语言的演变，古代经典跟学习者当时的语言已经有很大差距，因而读懂古代经典是学习的关键，而阅读古代经典首先面临的是会认、会写记录经典的汉字。这样，汉字教学就成了古代语文教学的重要的和首要的任务。由于汉字是形音义结合的语素文字或表意文字，字形本身不标音，加上字数多、笔画复杂、形似字多、异体字多，存在大量的多音字、多义字、音近字、同音字、同义字、近义字，使得汉字具有难认、难记、难写、难用等特点，尤其对初学者来说在识记、理解、读写上都有相当的困难。

为了解决识字教学中的难点，古人设计并实施了许多行之有效的教学方法，取得了许多可贵的经验，尤其在识字教材的编写上取得了不少的成就。由于汉字是语素文字，一个汉字一个音节，且古汉语中词以单音节为主，这样，字、词、音节就有相对的对应性。古人利用这些特点，采取几个字(3 个字或 4 个字不等)一组的整齐押韵的办法来编写儿童识字教材，这些教材因为整齐押韵，便于朗读记诵，因而很受欢迎。古代童蒙识字课本有西汉史游的《急就篇》，南朝周兴嗣的《千字文》，宋初的《百家姓》，相传为南宋王应麟编写的《三字经》，明程登吉编写的《幼学琼林》等，直到元明以后的各

种识字课本都采用了这种方法来编写。这些童蒙识字教材在教儿童识字的同时，还附带教一些知识，如姓氏、道德规范等，如《千字文》的“天地玄黄，宇宙洪荒”，《百家姓》的“赵钱孙李，周吴郑王，冯陈褚卫，蒋沈韩杨”，《三字经》“人之初，性本善，性相近，习相远，苟不教，性乃迁，教之道，贵以专”。《千字文》、《百家姓》、《三字经》等从宋代起形成了互相配合的整套童蒙识字教材，通称“三百千”，这些童蒙识字课本在全国范围内使用了 1 000 多年，共收汉字 2 000 多个，这些汉字的学习为初步阅读打下了良好的基础。

识字课本除了“三百千”外，还有一些所谓的“杂字”。“杂字”是流行于一时一地并带有浓厚地域色彩和行业色彩（如农业、商业、手工业等）的识字教材，“杂字”的用字用语比较通俗，甚至十分接近口语。杂字的起源不晚于“三百千”，品种很多，但由于主要流行于下层社会，不被重视，大多遗失，流传下来的大约有二三十种。“杂字”中还有一种看图识字的识字课本，现存有两种，一是《对相四言杂字》（刻于明太祖洪武辛亥年，即 1371 年），一种是《新编对相四言》，比《对相四言杂字》稍晚，内容小有修改①。这也许是世界上最早的看图识字教材。

2. 书法教学

识字教学和书法教学往往是联系在一起的。在学写字的入门阶段，写字和识字往往分开，这是因为汉字较为复杂，识字之初的字有的较为复杂，如《百家姓》开头的几个字“赵钱孙李”，不便于童蒙进行写字练习，写字练习之初往往选择笔画简单的“上、大、人”等字，而后逐渐复杂，就跟识字教学同步了。由于汉字结构复杂，儿童初学写字有一定的困难，为了便于儿童练习写字，古人设计了一整套的写字训练或书法练习的方法，如先写中楷（大约 5 厘米见方），以掌握笔画的安排组合，然后再学写小楷和大楷，有序进行。练习分为不同步骤，开始是“描红”，把印好的红字按照老师指点的笔顺描成黑的；第二步是“影写”，用比较薄的纸铺在印好的或者老师写好的字样上来摹写；第三步是写“米字格”或“圆格”纸，便于初学者安排好笔画和字的结构框架；第四步是在白纸上印好的方格里写；再后就是脱格写，而且从脱格写开始临摹名家碑帖，一般先临摹楷书，进而临摹行书或草书，甚至隶书和篆书，进入学习书法艺术阶段。汉字不仅是一种书写符号，而且是一种艺术，

① 从《对相四言杂字》和《新编对相四言》的内容来看，这类书产生于南宋，后代有略有增删的多种复刊本，现存的是其中的两本。

写字不仅要写得正确，写得规范，而且要写得好看，写成艺术。古人是把练习写字和练习书法艺术结合在一起的，由于采用了循序渐进的方法和合理的训练程序、手段，使得古人的书法教学成效显著。

3. 读书教学

由于记录汉语的文字是语素文字或表意文字（这在古代尤其明显），因而不认识一定量的汉字是无法读书的。这就使得古人的汉语文教学的识字教学和读书教学不能同步进行，一般是在识字教学取得一定进展之后才开始进入读书阶段。古人所读的书，一是讲究文道统一，以儒家经典为主，二是为了科举应试所必须读的书。从所读的书来看，首先要读的是“四书”（依次是《大学》、《中庸》、《论语》、《孟子》），这是儒家最为经典的著作。其次是“五经”（即《诗经》、《尚书》、《周易》、《礼记》、《春秋》）。再是其他一些经书，如《周礼》、《仪礼》、《尔雅》等，还可以读一些史书（如《史记》、《汉书》等）和子书（如《韩非子》、《荀子》、《吕氏春秋》等）。“五经”而下也是应试科举所需要读的书。

如果不准备科举应试，另有一些读物供儿童和少年阅读。就内容来看，一类是按照儒家思想进行伦理、道德、待人接物等行为规范的教育，如相传为先秦管子编的《弟子职》，东汉班昭编的《女论语》，宋代的《性理字训》、《小学韵语》、《小儿语》等，明清人编的《惜时贤文》、《弟子规》等。另一类是有关历史知识的读物，如宋代黄继善编的《史学提要》、明代赵南星编的《史韵》、清代王士云编的《鉴略》等。再一类是讲各科知识或各种掌故趣闻、名言隽语的读物，如唐代的《兔园策》和《蒙求》，宋代的《名物蒙求》，明代的《龙文鞭影》，清代的《幼学琼林》等。这些阅读内容反映了古人把道德、知识融合在一起的儿童教育理念。

文学教育尤其是诗歌教育非常受古人的重视，孔子就曾说过不读诗（《诗经》）便不会说话，后人更把诗歌创作看作文人的基本能力和素质。因而选定教儿童读诗、学诗的诗选在古代印行了许多，如唐代胡曾编选的《咏史诗》，宋代编选的《千家诗》等。《千家诗》选编了唐宋名家的优秀小诗，便于儿童诵咏，长期广泛流行，与“三百千”并行为“三百千千”。古人读书，从小便接受诗歌欣赏和创作教育，以致历代诗人不绝，使得中国成为诗的国度，古人留给后人许许多多的优秀诗篇，并出现了许多伟大的诗人。除了诗歌教材外，古人还编选了一些散文（文言文）故事作为儿童阅读材料，如明代以后印行的有《日记故事》、《书言故事》、《黄眉故事》、《童蒙观鉴》等。

就读书方法来看，古人诵读和讲读是分阶段进行的，一般是先识字，后

诵读，再讲解。而实际上，讲解是不被重视的，古人对文章作品的理解讲究的是“书读千遍，其义自现”或“神而明之”，诚如胡适所批评的：“从前教作文的人大概都不懂文法，他们改文章全无标准，只靠机械的读下去，读得顺口便是，不顺口便不是，总讲不出为什么要这样做，为什么不可那样做。”①这从一个方面说明古人在语文教学中对讲读是不重视的，强调的是诵读。诵读可以逐渐领悟文章的内容和写作技巧，并能逐渐培养语感，但毕竟费时耗力。另外，古人的读书教学过于强调对文章所包含的思想道德内容的掌握，有忽视语文作为工具的倾向。

4. 作文教学

为了科举应试，儿童在读完《四书》之后，便要开始学习写作，即学做八股文和试帖诗(学做试帖诗之前先要学对对子)。为了学做八股文和试帖诗，也有各种选本供学习者仿效，也有印行的历科考中举人、进士的考卷，即“闱墨”，供应考者参考模仿。除了学做八股文和试帖诗外，年龄稍大的学生还要学做“古文”和其他诗体。要做好八股文、试帖诗以及古文和其他诗体，必须学习各种范文，因而历代印行了许多诗文选本，如《昭明文选》(南朝梁萧统编选)、《玉台新咏》(南朝陈徐陵编选)、《文章正宗》(宋真德秀编选)、《文章轨范》(宋谢枋德编选)、《古文关键》(宋吕祖谦编选)、《文苑英华》(宋李昉、宋白等编选)等，其中诗歌选本影响最大的是《唐诗三百首》(清孙洙编选)，古文选本影响最大的是《古文观止》(清吴楚材、吴调侯编选)。各种选本大都有或繁或简的注释，也有解析和评点。可见，古人的作文教学是非常重视范文学习，模仿范文是学做作文的主要方法。教学中老师对范文进行串讲，学生则要熟读成诵，逐渐领悟其中的技巧。

秦汉至清末的传统语文教育，历经近 2 000 年的发展逐渐形成固定的范式，如语文教学的内容以儒家经典为主，强调道德、伦理在语文教学中的地位，不仅是教会学习者阅读理解经典，而且要接受经典的思想内容，因而传统语文教学教授的语言是已经在口语中不用的书面语；再如，传统语文教学的目的是明确的，主要是为了科举应试，科举考试自隋炀帝大业二年(606年)始，历时 1 300 多年，而科举考试的内容主要是儒家的经典，如唐代为《五经正义》，宋朝尤其重视经义论策，明清更是以“四书”、“五经”的文句为题，解释则以朱熹的《四书章句》等为依据；从教学的方法看，传统语文教学

① 胡适《中学国文的教授》，《新青年》1920 年 9 月 1 日 8 卷 1 号。

重视读、写，学习的主要方式是在背诵的基础上模仿。因而，传统语文教学不能很好地培养学习者听说读写的全面能力，也不能很好地培养学习者面向现实生活的实际的语文交际能力。同时更主要的是传统语文教学在学校教育中并不是一门独立的学科，而是经学的附庸，是和经学、史学、哲学、伦理学等结合在一起讲授的。所以严格地讲我国古代没有纯粹的语文课堂教学，传统语文教学的弊端是显而易见的。不过，传统语文教学重视阅读、写作，尤其是重视对经典范文的诵读和模仿，这有利于培养学习者的语感，值得今后的语文教学借鉴。

### 三、变革时期

鸦片战争之后，中国社会面临深重的社会危机，为了救国救民，大批有识之士开始学习和借鉴西方的学说，一些人看到中西教育方式、教育效果的差异，看到中国传统语文教学的弊端，从而开始中国语文教学的变革。由于资产阶级改良派和洋务派的推动，清朝末年（1905 年）废除了科举考试。1904 年清政府颁布了《奏定学堂章程》，在全国范围内兴办新式学堂，从此，中国教育才开始真正的分科、分班、分年级教学，这样，语文教学才可能成为一门独立的课程，如《奏定学堂章程》中规定的小学教育中就有“读经讲经”和“中国文学”两门课跟语文有关。于此前后，开始有了真正意义上的语文教科书，1903 年商务印书馆编印出版了我国第一套小学用最新教科书，其中《初等小学国文教科书》和《高等小学国文教科书》分 9 年 18 册，这对语文教学影响至深。稍后，商务印书馆又于 1908 年编印了《中学国文教科书》。这个时期还出现了专门研究国文“教授法”的文章。不过，语文教学的真正变革却是 1919 年五四新文化运动前后，五四运动使得传统语文教学受到深刻的冲击。这种冲击及其引起的变革主要表现在如下几个方面：

**1. 中小学普遍开设国语课**

五四新文化运动使得白话文成为正宗的文学语言，这样，五四以后小学里普遍开设了“国语”课，全用白话文，这就打破了文言文（国文）的一统天下。中学里虽然还开设国文课（文言文），但国语课（白话文）的比例越来越大。1923 年颁布了《小学国语课程纲要》、《初级中学国语课程纲要》、《高级中学公共必修科国语科学程纲要》等，这些纲要规定了国语课教学的主旨、标准、程序、教学方法等，使得国语教学有纲可循；同时还编印出版了《国语教科书》、《初级国语读本》、《高级国语教科书》等，这些教材选收了接近口语

的白话文作品,内容上也跟儿童的生活和学习的兴趣较为接近,如选用了一些中外儿童文学作品和用白话翻译的古代寓言故事等,并且随着学习阶段的提高,文学性强的名作家的白话文学作品也越来越多,不仅一些早期白话作品如《三国演义》、《水浒传》、《镜花缘》等的片段入选课文,而且更多的新文学作品如蔡元培、鲁迅、叶圣陶、冰心、郭沫若、徐志摩、茅盾、巴金、朱自清、老舍等著名作家的作品入选课文。从学习顺序看,国语课的开设就把识字教学和阅读教学结合起来了,一边识字一边阅读,在阅读中逐渐识字,使得读写能同步发展。

**2. 识字和阅读的工具得到改造,可阅读的书刊越来越丰富**

随着国语运动的发展,注音字母逐渐在语文教学中被使用,注音字母的推行大大方便了识字教学和阅读教学。随着字典、词典编纂的进步,供中小学生使用的工具书多了起来。另外,随着现代出版业的发展,供学生自由选择的阅读书刊如各种课外读物、报纸、期刊(如《儿童世界》、《学生杂志》、《中学生》、《新青年》、《小说月报》、《东方杂志》等)也多了起来,这些大大方便了学生的课外阅读,开拓了学生的视野。

**3. 语文教学理论和教学法的研究取得丰硕成果**

随着东西方现代教育理论的引进,许多学者开始探讨语文教学理论和教学法。像陶行知、夏丏尊、黎锦熙、叶圣陶、王森然、阮真、朱自清、浦江清等,都发表或出版过语文教学的理论和教学法方面的论著,如黎锦熙的《新著国语教学法》(我国第一部语文教学法专集)、王森然的《中学国文教学概要》、阮真的《中学国文教学法》、叶圣陶的《作文论》、夏丏尊和叶圣陶的《文心》、叶圣陶和朱自清的《国文教学》、蒋伯潜的《中学国文教学法》等。这些语文教学专家还在实践中运用过问题教学法、因文而异的程序教学法、设计教学法、道尔顿制教学法①、导学教学法、五段教学法、三段六步法、教学做合一法等。这些教学法研究虽各有不同的侧重,但共同的主张是:① 反对以阅读古代经典为主,提倡白话文,阅读跟社会生活接近的作品,包括文学作品和有实际应用价值的文章;② 反对注入式讲解和让学生呆读死记,提倡灵活科学的教学法;③ 提倡和鼓励学生广泛阅读,独立思考,学以致用。

---

① 道尔顿制是美国教育家伯克赫斯特于1920年创制的一种新的教育制度,因为创试工作是在美国的马萨诸塞州道尔顿市的道尔顿中学进行的,故名道尔顿制。

## 四、改革实验时期

1949年以后，在新的社会政治形势下，我国的语文教学进行了许多改革和实验。这表现在如下几个方面。

### 1. “语文”学科的建立和语文教学法的讨论

过去的“语文”课分别叫国文和国语，1950年中央人民政府出版总署编审局出版了全国统一的“语文”课本，即《初中语文》和《高中语文》，并且在教材的《编辑大意》中说明“说出来的是语言，写出来是文章，文章依据语言，‘语’和‘文’是分不开的”，就是说，“语文”是口头语言和书面语言的统一。“语文”的提出和界定使得语文学科得以正式确定。同时，这一时期关于语文教学法的研究也在前苏联教育理论影响下蓬勃开展。前苏联的教育学理论著作如《教育学》(凯洛夫著)、《语文教学法》(卡诺内庚著)、《文学教学法》(果鲁伯柯夫著)等被翻译到中国来。在前苏联教育理论影响下，语文教学界有针对性地开展了“关于《红领巾》观摩教学讨论”、“关于语文教学目的任务的讨论(文道之争)”、“怎样教好语文课的讨论”等，这些讨论的参与者有专家(包括前苏联的专家)，也有一线语文教师，推动了语文教学的深入发展。黎锦熙、叶圣陶、吕叔湘、张志公等学者都对新中国的语文教学法的发展起到了积极的作用，如黎锦熙在《新国文教学法》中提出“精读须‘手到’，熟读须‘口到’”等国文教学的一般原则；叶圣陶则强调口语和书面语并举、听说读写并重的语文教学观；吕叔湘也强调语文教学是口语和书面语的教学，要求语言和文字(书面语、文学)并举，以语言为门径，以文字为重点，达到语言和文字都提高的目的，吕叔湘还特别强调语言训练对思维训练的重要性；张志公强调“语文”的工具性，十分重视中小学语文教学中的口语训练。

### 2. 形成了几种语文教学的具体方法

由于语文学科的独立性和语文教学理论的发展，这一时期还出现了几种有价值的语文教学方法，如在业余学校、速成学校、扫盲学校推行的“速成识字法”，辽宁黑山北关小学的“集中识字”、“精讲多练”、“提早写作”的小学语文教改，上海市育才中学的“紧扣教材，边讲边练，新旧联系，因材施教”的十六字诀中学语文教改，北京市景山学校的“分批集中识字”教改等。识字教学的各种改革在《汉语拼音方案》公布以后都能充分利用汉语拼音帮助识字和读写。

### 3. 汉语和文学分科教学实验

这一时期语文教学的一次重大改革是汉语和文学的分科教学实验。

1953 年 12 月中央语文教学问题委员会向党中央提交了《关于改进中小学语文教学的报告》，分析了语言和文学混教的弊病，强调语言和文学必须实行分科教学。1954 年 2 月中央政治局扩大会议决定中学语文实行汉语和文学分科教学，与此同时，中央教育部陆续制订和发行了《初级中学汉语教学大纲(草案)》、《初级中学文学教学大纲(草案)》、《高级中学文学教学大纲(草案)》，并随后拟订了"暂拟汉语教学语法系统"，编出了《初级中学汉语课本》、《初级中学文学课本》、《高级中学文学课本》、《中学作文教学初步方案(草稿)》，以及相应的教学参考书。1955 年秋季分科教学开始试教；1956 年秋季开始，中学、中等师范学校的语文科分汉语、文学两科进行教学；1958 年停止。汉语和文学分科教学实验虽然时间不长，但对中学语文教学还是起到一定的促进作用的，一是文学课本选材广泛，民歌、民间故事、童话、寓言、古典诗词、戏曲、小说、散文等中外文学的多种样式都进入课本或课外读物，开拓了学生的视野，丰富了学生的知识，提高了学生的文学素养。汉语课也对培养学生系统的语言能力很有帮助，汉语课本及其相关文献还推动了汉语语法研究的发展。同时，通过分科教学，大大促进了教师业务水平的提高，推动了一线教师进行教学改革的积极性。但分科教学在大纲、教材编写，教学过程中都有一些缺点，如文学教材分量过重，文学跟写作的关系没能处理好，学生实际写作水平很难提高，教材分量太重，学生无法消化，汉语课过分重视语法知识，忽略了趣味性和实用性。

4.《全日制中学语文教学大纲(草案)》的颁布

1963 年 5 月教育部制订了《全日制中学语文教学大纲(草案)》，该大纲明确提出语文学科的工具性质，明确了语文教学的目的是"教学生能够正确地理解和运用祖国的语言文字，使他们具有现代语文的阅读能力和写作能力，具有初步阅读文言文的能力"，在教学方法上提倡精讲多练，启发式教学。这些提法都比较正确地反映了语文学科自身的特点和规律，对过分重视思想性的"极左"做法是一个反正。

### 五、语文教学的繁荣发展和思考

1978 年以后，语文教学进入了拨乱反正之后繁荣发展的新时期，具体表现在以下几方面。

1. 语文教学研究蓬勃开展

新时期伊始，叶圣陶、吕叔湘、张志公等老一辈学者就大声疾呼开展语

文教学的调查、研究、改革。语文教学专家和一线语文教师开展了一系列的语文教学的研究和语文教学的改革。1979年全国性的中学语文教学研究会成立,1980年小学语文教学研究会和语文教学法研究会成立,这些研究会吸引了大批专家和语文教师,推动了语文教学研究的发展。同时,刊登语文教学研究的专业杂志也越来越多,如《中学语文教学》、《小学语文教学》、《语文学习》、《语文教学通讯》、《中学语文》、《语文月刊》、《语文教学与研究》等。语文教学法(中学语文教材教法)作为学科成为师范院校的中文专业必修课程,并设有硕士学位和博士学位(隶属于"学科教学论"学科),培养了许多专业人才,朱绍禹、张鸿苓、陶本一等语文教学专家出版了数十种以语文教学法为名的著作,使得语文教学法作为一门学科更加成熟。

2. 语文教改实验逐渐形成多样化的态势

在教育学理论和语文教学法理论的指导下,广大语文教师在自己的教学实践中也创造出了许多具体的教学方法,如陆继椿的"分类集中分阶段进行语言训练"实验,于漪的"苦、得、趣"教学方法,钱梦龙的"三主"(学生为主体,教学为主导,训练为主线)、"四式"(自读式、教读式、练习式、复读式)的语文导读法,魏书生的"六步课堂教学法"(定向—自学—讨论—答疑—自测—自结),张孝纯的"大语文教育"观,章熊的"发展语言的同时发展思维"的实验研究,黑龙江省小学语文"注音识字,提前读写"实验等等。这些教学改革大大推动了语文教学的深入发展。

3. 语文教学大纲和教材建设

为规范语文教学,新时期公布了多个语文教学大纲,1978年教育部颁发了《全日制十年制小学语文教学大纲(试行草案)》和《全日制十年制中学语文教学大纲(试行草案)》。这两个大纲总结了过去中小学语文教学的经验和教训,重新明确了语文学科的性质、目的和任务,提出了语文教学在教材内容和教学方法上的原则要求,如小学语文教学的目的和要求是"培养学生识字、看书、作文的能力,初步培养准确、鲜明、生动的文风",中学语文教学大纲中首次强调要对学生进行严格的读写训练。从1986年开始,中小学语文教学大纲的建设进入了深化改革的新的探索时期,即围绕实施九年制义务教育这个中心,研究小学阶段和初中阶段语文教学的一些根本性问题。1988年国家教委出版了《九年义务教育小学语文教学大纲》和《九年义务教育初级中学语文教学大纲》两个"初审稿",1992年正式出版了九年义务教育《小学语文教学大纲》和《初级中学语文教学大纲》(试用本)。这两份大纲

较好地解决了语文教学中的一些问题：一是强调了听说读写 4 项基本的语文能力相互联系、全面训练；二是注意在语文能力训练中学习知识、锻炼意志、培养感情、发展思维，使智力因素和非智力因素得到协调发展，体现了义务教育注重素质培养的精神；三是使教学目标和教学内容进一步具体化、明确化；四是重视语文学科与其他学科、与实际生活的联系，注重语文的课外活动，注重语文作为工具的实际应用；五是适当减轻了难度，教学内容安排上有一定的弹性。

在大纲的指导下，不仅编写出全国性的语文教材（如人民教育出版社出版了多套教材），而且不少省市也编写出不同系列的语文教材，形成了在统一要求下的语文教材多样化的态势。

**4. 关于语文教学的几点思考**

总之，近 20 年来，我国的语文教学取得了许多成就，但也存在一些问题，这些问题的产生，除了一些学校或教师认为语文教学对提高中考和高考成绩（分数）实效不大而减少语文课时这种认识偏差的因素外，更主要的是一些理论问题没有很好地解决。这些问题包括：

(1) 关于语文学科的性质和任务

目前关于语文学科的性质有四种看法，即工具论、人文论、素质论、语感论。语文学科的性质是关乎语文学科体系和整个语文教学发展的关键，影响到大纲设计和教材编写，乃至课堂教学。

(2) 关于语文教学内容的优化选择

语文课究竟应该包括哪些内容，如何优化选择，对已经选择的内容如何排列，学生学习的顺序是什么样的，这些问题都很少经过科学的研究和实验，目前教学内容的选择和排列多是凭经验和感觉，缺乏理论的分析和论证，而解决这些问题需要多学科的合作。不解决这些问题就可能使教学内容的范围、深度、难度等没有一个明确的标准和要求，容易出现随意性和盲目性。

(3) 关于语文知识和语文能力训练的关系

这包含两个具体问题，一是语文知识对能力的培养有没有用，二是语文能力训练需要什么样的语文知识。曾经有一段时间，语文教学中，语文知识如语法知识讲得过多过深，效果不好；近来又有人完全否认语文知识的教学，如提出“淡化语法”，甚至要“扭断语法的脖子”。实际上语文知识的学习对语文能力的培养是很有作用的，问题是需要教哪些语文知识，如何教这些语文知识，如何处理好知识和能力的关系，人们提出的“精要、好懂、管用”应

该是一个很好的原则。

(4) 关于文和道的关系

过分突出和强调语文课的思想性或政治性,固然是片面的,完全否认语文课的思想性同样也是片面的,那么如何协调好这两者的关系呢? 如何在具体课文教学中处理好这两者的关系呢? 另外,如何历史地看待一些课文中所体现出的思想性,以及这些课文中的思想内容跟现实社会的关系如何等等,都是需要认真思考的。

(5) 关于教和学的关系

出于语文教师的责任感,老师总希望教给学生更多更深的知识,而不考虑学生的兴趣、理解能力和接受能力,这就使得教和学产生矛盾,这一矛盾直接影响了语文教学的效果。今后的语文课堂教学应该理顺教和学两方的关系,使教学双方处于和谐互动的氛围。

(6) 关于传统语文教育的研究和民族化、现代化语文教育体系的构建

语文教学既要继承中国语文教学的传统,吸收传统语文教学的精华,又要使语文教学现代化,做到"回头看"和"向前看"的结合。语文教学的现代化是指语文教育思想、教学内容和教学方法、手段的现代化和科学化,语文教学的民族化是继承和发扬传统语文教学的精华,突出汉语文的民族特点。只有把现代化和民族化有机地结合起来,语文教学才可能做到高质量、高效率和学以致用。

## 第三节 国外语言教学的发展

### 一、古代和中世纪

罗马帝国以前,罗马人把希腊语作为第二语言来学习。他们聘请希腊语私人教师,把说希腊语的奴隶叫到家里,或雇佣会讲希腊语者做奴仆,他们通过这种方法来学习希腊语。随着罗马帝国的扩张,其他民族也开始学习拉丁语,直至这种语言成为西方世界的国际语言、教会和国家的语言,并且在相当长的时间内这种语言也是学习的惟一语言和学校里惟一的教学媒介。而且,在欧洲的一些国家里,这种状况一直延续到现代。

因此,在欧洲,人们开始关注语言教学的教学法是同教授拉丁语相联系

的。在中世纪，拉丁语是教学语言，教学也基本上局限于教拉丁语语法，而这种语法是为了使教士们说、读、写第二语言而设计的。所有学科的教学也基本上都是用这一语言进行的。

## 二、文艺复兴时期

印刷术发明以后，有人对采用不合适的方法教授拉丁语提出不同意见。印刷机最初的一种用途，就是翻印希腊语和拉丁语古典著作，这些著作都是在欧洲广为传播并因而成为教授内容的。可是，这些拉丁语古典著作所使用的语言比当时欧洲学术界使用的语言要老了几个世纪。尽管如此，这些古典著作的语言仍被看作是拉丁语的原始的正确形式，看作拉丁语语法和教授这些语言的教学法的最基础形式。但以古典著作为基础写成的拉丁语语法逐渐变得更冗长、更复杂，结果拉丁语语法研究本身变成了目的，而不是为阅读这些古典著作所作的一种必要准备。而且，当时的法令规定，必须用这种古典的拉丁语代替当时所说的那种形式的语言，这渐渐把拉丁语变成了一种死的语言。最终的结果，使得拉丁语慢慢让位于欧洲各民族的语言。

在这种情况发生以前，也有过多次改进拉丁语教学的努力，希望革除为语法而学语法的作法。1532 年戴迪 · 马里尼斯(Marinis)出版了一部简明语法，目的在于“从学生中造就出拉丁语学家而不是语法学家”。卢瑟(Luther)和美兰克森(Melanchthon)也反对学习过多的形式语法和教授语法规则。

蒙田(Montaigne)曾在一篇论文中提到，他是在不学规则的情况下学会拉丁语的，因为他父亲要求家里的每个人，包括请来教该语言的德国籍私人教师，同他都只说拉丁语，对于孩子，更是在他们学习母语以前就让他们说拉丁语。可蒙田进入大学以后，原来说得非常流利的拉丁语很快开始遗忘。不过，在那时的大学里，他一开始很可能只学习语法规则。

在此时期，拉迪希乌斯(Ratichius)也抱怨当时盛行的那种以牺牲阅读为代价去强调语法的语言教学。不过，根据新的原则制定新的语言教学法这一工作直到他的后继者、捷克著名的教育家夸美纽斯(Jan Comenius)才完成。夸美纽斯在教授阅读和口语时，运用模仿、重复及大量的联系等方法，而不用规则。他认为语法是间接地通过归纳法学到的。1631 年，夸美纽斯出版了《语言入门》(Ianua Linguarum Reserata)一书，书中的词汇不多，只有几千词，每个词都有例句，简要说明其意义。第二年他又出版了《大教学论》(Didactica Magna)，这部著作奠定了现代教育学的基础。语言教学一

直是夸美纽斯最感兴趣的课题,他的《语言的最新教学方法》(Linguarum Methodus Novissma)是最早用归纳法教学语言的一个尝试。他的晚期著作《画中世界》(The Orbis Pictrus)更是第一部通过图画来教语言的著作。

## 三、17 世纪和 18 世纪

夸美纽斯去世后,他的多数著作很快被多数人遗忘,这种情况一直延续了两个多世纪。但是,仍有一些对语言有兴趣的人时时使用他的某些思想。洛克(John Locke)接受了夸美纽斯和蒙田的一些观点,还将它们写进他的《关于教育的若干想法》(Some Thoughts Concerning Education)一书中。

洛克认为:"语言不是由技艺的规则构成的,而是偶然形成的,是接受普遍用法的结果。一个能把语言说得很好的人除此以外是没有其他规则可用的。依赖记忆以及从潮流中学得的说潮流话的习惯,而不是信赖其他什么东西,是说话贴切的保证。换另一句话来归纳,也就是说话只能是死记硬背……"

巴则多(Basedow)也继承了夸美纽斯的观点,后来他又受到了卢梭(Rousseaw)自然教育思想的影响。不过当时的教育家反对他,他们认为他的思想危害了古典语言在教育中的地位。

到 18 世纪末,教授拉丁语法本身已经变成了目的。拉丁语已不再是教学媒介,教授和应用拉丁语规则变成了某种智力练习。在当时,有一些学校,引进少数几种现代语言,仍用教拉丁语的方法教,同时用作智力训练。

## 四、19 世纪

19 世纪初情况发生了很大的变化。雅科托(Jacotot)和图森(Toussaint)以詹姆斯· 汉密尔顿(James Homilton)为榜样,倡导人们回到老路上去,通过他们使用的第二语言的课本学习归纳性语法。但是,由于他们使用的课本不是文学作品就是《圣经》,这对于初学者来说,就太困难了。为了克服这个问题,塞登施迪克(Seidenstuck)在 1811 年成功编写一种课本,这一课本用简单句写成,却基本上包含了该语言的主要语法特征。后来阿恩(Ahn)和奥伦多夫(Ollendorf)继承了他的这一创新做法。不过,总的来看,塞登施迪克所选用的句子都是没有系统的,后来卡尔· 浦洛兹(Karl Plotz)从中总结出一条原则,这一原则很长时间内影响着欧洲的语言教学。浦洛兹的教学法主要包括两个部分:① 规则和范例;② 要翻译成第二语

言的句子和要翻译成第一语言的句子。在整个19世纪,学校的语言教学基本上都遵循了浦洛兹的方法。

19世纪后半叶,出现了许多逆潮流的事件,这些事件最后发展成了一种运动。这一运动的先驱者克洛德·马塞尔(Claude Marcel)在1867年发表的关于语言教学的论文中提出废除翻译和语法规则,提倡语言教学从理解课文和加强听的训练入手,接着是阅读简单而熟悉的材料,最后是说和写。1874年,索夫尔(Sauveur)主张在使用相似原则的同时,在教授学生能理解的语言的时候可以增加语法和翻译。到了1866年,埃内斯(Heness)开办了一所用"自然教学法"教语言的学校,目的就是对上述几个学者提出的语法翻译法作出激烈的反应。

19世纪80年代,语言教学中又增添了新的内容——体力活动。1880年,古安(Gouin)出版了《语言教学与学习的艺术》(Art d'énseigner et d'etudier les Langues)一书。该书当初在法国并没有得到重视,但其译本在其他国家如英国、德国和美国等却产生了很大的影响,也成为直接法的主要思想源泉,古安把正在发展的现代心理学科学的原则应用于语言学习,这些原则就是思想的联想、形象化、通过感觉学习、兴趣的中心、日常情境里熟悉的活动等。在他的教程里,每个句子都同其他句子相衔接,构成一种"选择",每种选择又同相关的"选择"相联系,共同构成一个"系列"。系列又同其他系列构成一个更大、更一般的系列。说出每个句子的同时必须表演。教程中全部词汇约8 000多个,这些词都与某一种活动有关。在古安的思想中,句子是言语的基本单位,每个句子必须与另一个句子相联系才可能构成一个链。这些句子链描写学生(而非教师)感兴趣的行为和活动。

与此同时,语言教学中又增添另一个新的内容——描写语音学。从19世纪中叶开始,就有很多学者研究过描写语音学,只有塞斯(Sayce)把它应用于现代语言教学,维埃托尔(Vietor)进一步把它归于某种教学法。维埃托尔在1882年出版了一本小册子,引起关注。在书中,他分析并抨击了浦洛兹学派的语法翻译法,提出语言教学的一个新路子。他与他的支持者以口语为出发点,详细设计了一种教学法。在这一教学法中,词不能被孤立地教和学,它必须在句子中使用,因为句子是言语的基本单位。句子之间不是无联系的,它必须与学生感兴趣内容的其他句子在一起,出现在上下文中。新的材料应通过姿势、图画和已知的词来教授。阅读的引入应该安排在较后的阶段,引入时必须对能获得外国及其文化知识的阅读材料进行分级。

所有的语法知识都必须通过课文的学习，采用归纳的方法获得。

这些观点在很长的时间内一直存在着争论，虽然如此，一开始在德国和斯堪的纳维亚半岛各国，后在法国、英国、美国得到了一些直接的支持，它与古安的某些原则相结合，产生了直接教学法运动。

由直接教学法产生的教学，一开始是很混乱的。教师替代了书本，而教师又常常并没有通过做动作进行教学的技巧，所以教师常常可以随心所欲，有时甚至请一个说本族语的人来作模仿或发音合作人。这一切实际上是在实践中把维埃托尔和古安的原则简单化处理了。所以有人将它们与“自然教学法”混同了。自然教学法认为语言必须以儿童习得母语的方式才能学会。学习者如果只能理解所听到的一点点，也没有关系，因为“自然选择”的原则最后会给他所需要的。即使他一开始说话时总是出错，他也可以在此过程中熟悉这种语言，这会为最终辨别并改正自己的错误打下基础。

基于上面提到的原因，很自然的直接教学法也不乏批评者。批评者们指出：儿童的自然选择理论与第二语言学习是不同的，学习过程中太多的困难会导致混乱、错误、浪费事件和发音不准。而且，这一教学法对教师提出了很高的要求，要求教师能熟练使用第二语言且有教授这种语言丰富的经验及教学过程中的应变能力。

这些批评也使得直接教学法的倡导者们最终承认有必要使方法系统化。口语所起的举足轻重的作用以及训练听力的必要性促使人们去努力实现发音的标准化，其结果是一开始就把系统的元音和辅音操练归入了教学法，并经常用语音音标作为辅助手段。掌握语音后，再教整个句子及其意义。

### 五、20世纪

到了20世纪初，直接教学法的教科书才开始遵循明确的模式。典型的课本都严格遵循不使用学生母语的原则，最先教口语，再教阅读，最后才练写作。

学习口语的第一步，就是借助音标来学习语音。这些语音在句子中使用，而这些句子的意义可以在教师姿势和才智的协助下，通过图画和定义来传授。语言结构通过推理和抽象获得。阅读材料也都用当代语体写成，内容一般是外国的生活与文化。

直接教学法在系统教授语言的结构及其意义方面也做过一些努力。教

材大体上分了等级,下一节课的内容基本上与上一节课的内容存在着必然联系。这使教师掌握特殊教学方法的必要性得到了认可,特殊的假期课程就是基于这样的考虑开设的。1898 年,在维也纳召开了一次现代语言教师国际大会,会议接受了这样一条原则:在组织每一节基础课时,采用对话或演剧等直接教学法的技术应占一定的时间。这就使得直接法运动进一步得到了充实。1900 年,在莱比锡召开的一次会议上,通过了一项在高年级应用新原则的动议。新原则的主要内容有:在课堂上一般只准使用第二语言,使用以现代散文、介绍外国文化的材料和现代文学作品为内容的课文、归纳语法、复述性的笔头练习等。

"直接教学法"作为术语的确立是 1901 年的事。这一年,法国公共教育部的一项通知中对这一术语进行了描述。到 1902 年,由于 Leygues(当时的部长法令)的颁布,直接教学法成了法国外语教学中惟一被正式批准的教学法。同年,这一教学法与埃维托尔的语音学在德国也成了法定的教学法。不过,虽然沃特·里普曼也将它介绍到了英国,却没有在那里成为法定的教学法。

采用直接教学法的国家在教学过程中也出现了很多困难。从事教学工作的教师中,只有少部分受过使用这一教学法的专业训练。许多教师达不到教学法所要求的使用第二语言的熟练程度,也缺乏技巧和能力。在技巧方面,人们迫切感到需要使用这类教学法获得成功的教师之间交流经验。语言教师们开始访问使用这一教学法并获得成功的一些中心,将观察到的技巧描写下来并予以发表,这对完善这一教学法起到了积极的作用。比如,1909 年,皮慈切尔(Pitschel)观察了一个学校正在实行的教学法,此后,他就如何有效地使用图表、壁画、广告和歌曲等方面发表了一系列研究成果及评论,影响较大。

在使用直接教学法的过程中,针对这一方法存在的问题,每个国家都根据实际情况采用了某种折中方案。在德国,经过修改后,它才比较适于传统的语言教学,最终该方案得到了归并,成为后来称为"综合教学法"的一部分。这一结合是直接法中采用的语音学、直觉法、归纳法、现代课本与传统法对语法作系统研究的结合。在法国,采用折中方案的结果是产生了"活用法"。在比利时,这一教学法一直得到梅洛(Melon)的支持,1895 年,政府颁布了法令,规定口语应作为一切初级语言教学的基础,并应通过自然的和直觉的教学技巧来讲授,不过,这一教学法最后也演变成了某种折中方案。

在英国,由于斯威特(Sweet)、里普曼(Ripman)和芬德利(Findlay)等

人的努力，支持者们积极使用这一方法并取得了成功。但是，当这一方法普及到一般的学校后，情况就变得很不如意，因为这对教师自身的要求很高，而且需要每一节课都花费大量的时间进行准备。而结果常常是教师花费了很多精力，结果却不能令人满意，最后，教师们又回到了某种形式的语法翻译法。

在美国，《十二人委员会的报告》简洁地表明了对直接教学法的看法。该委员会由12位著名的美国语言教师组成，任务是在语言教学的课程设置及教学法等方面为全国教育协会提出建议。该报告指出："不应该把培养谈话能力本身当作是最重要的事，而应该看作是取得语言学知识和文学修养这样一些更高目标的辅助手段。"在托马斯的主持下，该委员会概述了语法-翻译教学法，内容包括教授语法，翻译简单的短文，练习听和发外语的语音等。该委员会认为，外语教学的主要目标是使学习者一看见这种语言就能翻译，最终是为了阅读。这种方法将恰当使用本族语言放到了非常重要的地位。这一方法对合适的教科书提出了新的要求，为满足这些需求，1900年至1914年间美国先后出版了一些很好的教科书。

由于学校课程和教科书并未把培养用外语进行谈话等技能放在其中，但社会又有了实际需求，一批语言商人就承担起了这项工作。贝里慈(Berlitz)是其中比较著名的人物之一，他于1878年开办了一所语言学校，后来，他的家族将学校的数量增加到200多所，遍及全国各地。由于只能用口语进行教学，他们的教学法类似于直接教学法。

由于训练口语的需要，美国和加拿大的一些教师从欧洲引进了一些属于"革新"的新技巧。但结果并未如人所愿，因为这些新技巧把语法教学法所忽略的东西(即实际地正确使用口语)推到了另一个极端，它要求把掌握口语作为首要目标，而这对于中等程度的学校是难以达到的。这一教学法不仅要求教师熟练掌握外语，在精力和体力等方面的要求也特别高。对此，有关部门制定了一个大规模的规划，拟用试验来解决教学中存在的问题。在美国，1920年到1935年是一个积极试验的时期，在此期间，各类目标都得到了评价。在此基础上，许多计划得以指定，也取得了很多一致性意见。

二次大战期间，美国大量的陆军和海军驻扎在世界各个角落，但必要的外语人才非常缺乏。为了给这些远离国家的军队提供能流利地说多种语言的人，美国陆军在一些大学的支持下，根据陆军特别训练计划建立了自己的

学校。有些学校几个月后的结果给人们留下了深刻的印象，人们就相信这种培训中包含着使语言教学取得成功的奥秘，但事实上并不存在。其实为了取得教学的目的，在教学过程中并非使用了某种方法，而是采用了各种方法及技巧。

战时语言学校取得的成绩是惊人的，结果战后许多中小学甚至大学都试图走它们的路子。但实际上效果并不如意，因为学生不可能将全部的时间放在学习语言上。

回顾一下语言教学的发展，可以看到，它首先从古代和中世纪的积极使用拉丁语口语转到文艺复兴时期通过学习语法规则来学习语言。后来夸美纽斯回到了口语活动，浦洛兹又转到语法规则，直接法则再次突出了言语。不过，在发展过程中，从对古典语言的要求转到了对现代语言的要求。一直到 20 世纪中期，教学现代语言的技巧并没有什么变化。

## 六、教学方法的构成和类型

在上面的讨论中，一再提到了教学法，这个术语是什么意思呢？它是不是仅仅指教什么、怎么教和何时教呢？其实不同的人对教学法有不同的理解，对有些人来说，它可能指的是教学一组程序，对另外的人来说，则是不要教学程序。对有些人，它指语言技能是第一位的，对另一些人，则指的是词汇和结构的类型和数量等。

教学方法之间的区别仅仅在于作为一种教学法的教学法所必须包括的内容上。一种教学法必须包括一切教学所包括的东西，不管所教授的是语言还是数学，是艺术还是物理。

一切教学，都必须包括某种选择，某种分级，某种引出和某种重复。要选择，是因为教授整个领域的知识是不可能的；要分级，是因为不可能同时讲授所有内容，它们必须有先有后；要引出，是因为教学就是要讲清什么或试图讲清什么；要重复，是因为不可能一次就能学会所有的东西。

在过去几个世纪中发展起来的很多教学法，到现在仍然在全世界各个不同的地区使用着，使用得比较普遍的教学法有：① 直接教学法；② 自然教学法；③ 心理教学法；④ 语音教学法；⑤ 阅读教学法；⑥ 语法教学法；⑦ 翻译教学法；⑧ 语法-翻译教学法；⑨ 综合教学法；⑩ 单元教学法；⑪ 语言-控制教学法；⑫ 模仿-记忆教学法；⑬ 练习-理论教学法；⑭ 同源教学法；⑮ 双语教学法。

## 第四节 语 言 理 论

语言理论对语言教学有什么现实意义呢？每种语言教学法的产生、应用和教授归根到底都取决于教学法倡导者或教师对语言所作的假设。比方说一种教学方法建立在语言是一整套词汇的假设之上，而另外一种却建立在认为语言是一个系统的假设之上，两种教学方法就一定有很大区别。再如认为语言是传统语法还是现行用法的差别，也会导致教学法的不同。方法论可能产生于一种无意识的假设，但语言教学方法已变得越来越科学，因此，它总是越明确地建立在语言理论假设上。

许多知识领域都涉及到语言，而且有些理论也详细描写了语言的规律。不同知识领域涉及到不同的事物，或者用不同的方法研究同一事物。从各个不同的角度看语言也是很正常的。如对哲学家，语言是思维的工具；对社会学家，语言是一种行为；对逻辑学家，语言是计算模式；对统计者，语言是有意无意的挑选；对语言学家，语言是一套任意的符号系统。

与中世纪以前语言理论比较，现代语言理论更关心语言怎样发挥作用，因此现代语言理论的各项原理大都根据对一种或多种语言的观察而得出的；相应的，一种理论到底如何取决于它观察的是什么或怎么观察的，必须根据语言理论的主要特征对理论进行比较，才能确定出它们的差异和相似点。理论主要涉及概念是否妥当、语言的本质、语言的诸方面、术语等。研究这些方面，就可以得出语言理论的相互区别，以及它们对语言教学方法可能产生的影响。

### 一、概念的妥当性

在语言理论的发展过程中，它常常是与哲学问题联系在一起的。如罗素认为大多数哲学问题都可以归结为语言问题。维特根施坦的哲学分析大部分是对日常语言的分析和对词的功能的研究。卡西雷把语言看作是一种像科学思维和宗教思维一样独立的心智活动。

哲学家和语言学家所关心的东西各不相同。哲学家关心的是直接或间接地证明语言学的陈述，语言学家所分析的许多语言陈述是与逻辑没有关

系的。语言学家关心的是在某种语言中一切可能的陈述形式和意义，如疑问、命令、对价值的判断等。不同的假设对语言基本概念具有重大意义。

1. 语言和人的概念

既然语言是人的一种活动，那么人的活动还包括什么呢？对此，存在着不同的看法，从而产生不同的语言观。有人认为人的活动可以看作是完全的身体活动，还有的看作是心理活动。

前者是机械论观点，它认为，人的精神是身体的延伸，所不同的只是精神活动无法观察。身体与心灵、有生命与无生命在本质上是相同的，只不过存在着程度上的差异。人的所有活动，包括语言，都是一连串的物质因果序列。这种因果序列可以用实验所提供的证据来进行研究。因为，人对语言刺激的反应，在本质上与动物对周边环境的物理反应是同质的。由于语言所表达的意思无法直接观察到，持机械论观点者自然会忽视它们。

因此，持这种观点的语言描写以及由此而形成的语言教学法常常把语言看作一套形式体系，而不是无数意义的集合体。布龙菲尔德就是持这一观点的典型人物之一。

将人的活动看作心理活动的是心灵主义观点，它认为语言行为主要是心理行为，虽然语言与说话时的物理行为有密切关系，但却是不同类型的。它们并非程度上的差别，而是类别差异，因此不能将语言归为物理活动。

心灵主义者还认为，人类语言不同于动物的行为。动物可以通过训练形成条件反射。人的许多行为都是自觉行为，它和动物的条件反射有本质的不同，语言作为人类和社会的现象，不能简单地看作物理或动物的行为。

根据心灵主义观点来描写语言和语言教学法，就常常特别重视意义，而非仅仅重视物理形式。索绪尔是这一观点的代表人物之一。

2. 语言和知识的概念

一种理论，或需要通过对事实的观察和分类来描写语言，或需要通过直觉和构筑模式，从而把一切可能的事实演绎出来，前者是归纳法，后者是演绎法。

归纳法认为对语言惟一有效的陈述方法就是观察语言事实，将它们进行分类。语言学家常常收集言语行为的样品，观察它们并对其不同点进行分类。这一方法虽然只能分析一种语言全部言语行为的一小部分，但能对所观察的现象进行概括，并将概括的结论泛化到其他未观察到的部分。

按照这种方法来描述语言和语言教学方法，特别重视亲自观察、易于分

类的语言特征。哈里斯所使用的就是这一方法。

演绎法研究语言理论依据的就是理论科学，研究者发现了某个模式，构拟出一个理论模型，再试着从中演绎出多少东西来。推导出正确模式要靠科学直感，他必须对研究的对象非常了解，所以一个使用演绎法的语言学家首先必须掌握他所描写的语言。

按照这种理论来描绘语言或进行语言教学，常把重点放在语言的最大模式，也就是词类体系和句法体系上。

3. 语言的本质

语言理论的建立可以基于以下假设：语言从本质上说是：① 实体；② 形式；③ 兼有实体和形式。

语言可以看作有可看到、听到、感觉到、想像到的事物构成的实体。当然我们所想到的事物的实体与谈话或书写的实体（语音、书写符号等）是不同的。虽然现代语言理论已经将词与其所表示的事物区分开来，但人们还是常常将它们看作相等的东西。有些语言理论也仅仅把所看到或听到的语言进行物质的研究，如只研究表达语言的实体——语音等，这些理论大多是语音理论。有些语音学家坚持认为，语音实体是语言研究最根本的东西。

但 20 世纪前半叶最有影响的语言学理论，并不把语言看作实体，而是看作形式。有人曾经把语言完全看作思想的形式化，认为语言不是思想本身，而是独立的符号形式，这种理论不能说明语言的声音和形式，也不能说明实体和表达与声音的联系。不过把语言看作表达手段形式化的理论，不仅已经用于描写语言，也已经用于语言教学法。

布龙菲尔德和索绪尔都把语言看作一种形式，而不是一种实体，都把形式差别看作语言的主要特征。布龙菲尔德认为语言存在的惟一必要条件是，每个音位应与其他一切音位有不可混淆的差异。索绪尔认为构成语言实体的模式必须是任意的。

许多重要理论在把语言看作形式的同时，也坚持认为语言就其本质而言是个实体。这一理论认为语言可以是它所代表的思想，也可以是表达它们的语音，当然可以两个方面都是。

萨丕尔把语言看是自足的、创造性的符号组织；语言不仅指称不借助于语言而获得的经验，也实际上规定着我们的经验，因为语言的结构极其完整。沃尔夫把这种观点详细阐述为理论，认为语言的结构或语法不仅是表达思想的复制工具，它本身就是心智活动的大纲和指南，可用来分析各种印

象，综合头脑中的一切东西。

弗思(Firth)理论认为，作为人类特有活动的语言，既不应该把使用语言的人们的心智、思维和思想排除在外，也不应该把使用语言的情境排除出去。弗思的学生发展了他的理论，如韩礼德(Halliday)指出，虽然语言的基础在语境中，但是语言理论并不包括对语境的分析，语言理论关心的只是与情境有关的语言形式。这种语境用语言的语法和词汇来表达，借助语言的音位系统和正字法，用实际的语音和常用的书写手段来表达。

## 二、语言的诸方面

每一种语言理论，都可以从 3 个不同的方面来观察语言：① 语言听上去和看上去是什么样的；② 语言是怎样工作的；③ 语言是如何发展的。

1. *语言作为一种状态*

如果研究时注意的是语言听上去或看上去是什么样的，这种理论就是把语言看作某种固定的状态。这也可以从两个方面来看。

1) 就从属关系来看，语言可以被看作是人类的思维或人类的行为。在 20 世纪初期，布律诺(Brunot)把语言划分为 5 个思维范畴：存在、事实、环境、情态、关系，然后试图说明在法语中这 5 种关系是如何表达的。

也有许多语言学家把语言看作是由许多行为单位组成的。如派克(Pike)把语言放在人类总体行为的情境中，包括语言在内的全部社会活动可以分成许多意义单位，它们又包括若干没有意义的变体。另一位语言学家齐夫(Zipf)也把语言看作行为，他认为语言状态从属于省力原则。依据这一原则，人类行为中使用的语言形式是一种妥协，说话者力图通过最小限度的努力传达自己的思想感情，听话者同样如此。

2) 把语言看作一种独立状态。索绪尔把语言作为一种结构，他的思想已经发展成为结构主义理论。语言符号论是其中之一，这一理论把语言作为内在关系系统来考察，作为只包括关系或功能的自我存在的整体。布龙菲尔德和乔姆斯基也将语言作为结构来考察。布龙菲尔德认为结构是语言中最小单位的线形组合，即语言音位的一个序列。乔姆斯基则认为语言结构本身就是一种理论，它能生成一切句子，且只能生成符合语法的句子。

将语言作为结构的观点，也是语言类型学研究的理论基础。类型学不考虑传统的语系分类，也不考虑语言的来源，它认为渊源不同的语言，也可能属于同一种结构类型。

将语言看作独立状态的另一种观点是将语言看作一种交际系统。虽然交际不是语言，但语言却能用于交际。从这一观点出发，语言可用于传输信息，从代码库中选择出符号系列。实际上我们使用语言，就好像在选择一个接一个的词。而一旦词被选定，紧接着作出的选择要受概率法则支配。

通讯科学研究的是中性语言。这就需要对语言作大量的统计分析，然后用专门的方法研究统计结果。统计分析对研究语言教学法有用，因为它能提供最常使用的语言项目的信息。

2. *语言是一种活动*

看语言是怎样工作的，就是把语言看作一种活动。这一观点关心语言怎样起作用，或如何使语言发挥作用。它也有两种情况。

1）把语言看作一种思维活动。这又有两类：第一类是将语言看作心理活动。它的基本原理是，使用语言引起的心理活动必然要用一定的时间。研究的任务是辨认这些心理活动，将它们转化为心理时间，从而展示语言活动的心理过程。这种研究一般最初要限定抽象程度，因为人们的思维正是通过抽象来捕捉和表达经验世界的。第二类是把语言看作是对所感知的东西的一种直接的、类似于动物的反应。语言好像成了一系列的条件反射。从这一观点出发，教授和操练的是不加分析的语言单位，也就是经常与特定情境相联系的完整话语。

语境刺激-反应的观点则与此不同，它把语言看作在刺激中的常在变体，以适应语言反应的相应变体，引导学生观察情境中每个成分和与之相对应的反应成分的关系，使他们抽象出语言模式。

2）把语言看作大脑的活动。这一观点将人脑当作一种物质活动来研究。现在主要有两种方法进行研究。

第一种方法是通过大脑的局部障碍来分析言语反应。既然所有的刺激物都在大脑里留下印迹，而且一个语言符号就是两种刺激物——声音形象和概念间的一种联想，那么在理论上这个符号就可能在大脑中找到。研究人脑语言活动的绝大多数人已不再相信语言属于心理形象。语言活动更像是交错于网络之间的神经脉冲。近年来，神经学家引导语言学理论的探索向着不同方向发展。人们已经发现，大脑的某些失调与语言的某些特征之间有相似关系。雅可布逊(Jakobson)曾把丧失记忆的基本类型与语言的基本特征作比较，有一种记忆丧失影响到把词按正确方式组合在一起的能力，

还有一种则影响到用一个单位来替换另一个单位的能力。

第二种方法是建立人脑模型、理论模型和工作模型。建立理论模型的目的在于把语言作为大脑的一种活动来研究。其中有一种类似色谱层析器，也就是一种语言假设的生成机械装置。使用时，把词和概念都看作"经验分子"，而"意义"粒子则是构成经验分子的原子。意义的原子则被看作能激发其他原子活跃的"信息束"。

工作模型就是所谓的"电脑"，目的是用它来了解语言在人脑中是怎样起作用的。

设计语言翻译机的努力已经大大推动对大脑语言活动的探索。不过，有些数学家不认为语言翻译机与人脑结构有相似之处。

**3. 语言是变化的**

有些学者把语言作为状态或活动来考察，不过也有人把语言看作某种持续的变化来考察。

(1) 时间上的变化

就每个人而言，语言在不同的时间内表现出的特点是不完全相同的。在社会方面，不同时期的语言有不同的特点。分析语言的各种时间变体的学问叫历史语言学。现在，一般把语言看作永远处在变化中的一种代码，这种变化不是为了适应某种特定的目的发明出来的，而是从成千上万人的多种需要的相互关系中产生出来的。语言一直在变化，而且现在这种变化仍在继续进行。

(2) 空间上的变化

语言的空间变体也产生了各种不同的语言理论。语言学家已经研究过某种现代语言的空间变体，也研究了可溯源至共同母语的语族的空间变体。前一种研究就是区域语言学，后一种研究就是比较语言学。

区域语言学产生的理论可以解释在操某一语言地区为何不同的地方存在着语言变化。这种变化不仅表现在语音方面，也表现在词汇和语法方面。区域语言学可以为语言教学提供经过证实的用法，从而可以决定教授何种形式。它可以帮助教师和教学法专家区分哪些形式是区域性的，哪些用法是普遍性的。

比较语言学得益于19世纪对进化论的研究，得益于能证实各生物体的起源及其相互关系所需要的技术。关于亲属语言起源和扩散的谱系理论在语言学者对比较语言学的研究中得到发展和改进。

## 三、术语

不同的语言理论还存在着一个明显的区别，那就是所使用或所创造的讨论语言问题的词语有差别。

### 1. 语言不同

语言学理论不是用某一种语言写成的。由于所用的语言不同，用来描写各种概念的词语数量就可能存在着差异。即使词语相同，它所表达的意义范围也不可能完全相同，如英语中有 language(语言)和 speech(言语)，可法语中有 langue(语言)、langage(语言)和 parole(言语)，它们所表示的意义不可能完全对等。

这一情况就使得有些理论学家提出假设，某一理论如果最初换一种语言创造，理论本身可能就会有所不同。

### 2. 术语不同

理论家肯定要使用一些术语。有些理论家为了描述的方便和准确性等，发明一些新的词语，这些词语大都被接受，并被收入相关的语言学词典。

不过，不同理论使用的术语不可能总是与新概念永久相适应，对此，语言学家常常是创造出一些新词语。

不同的语言学流派和语言学理论都创造出各种成套的术语。即使同一流派当中，不同的作者也可能有自己特定的术语。术语的不同也常常能反映出理论的差异。

### 3. 意义不同

同一流派的作者可以使用同一个术语，但不能保证术语的意义完全相同。例如索绪尔和纪尧姆都使用 signifiant(能指)、signifiē(所指)和 signe(符号)，但意义却不相同，纪尧姆的 signifiant 与索绪尔的 signe 相同，而他的 signe 则带有索绪尔的 signifiant 的意义。

这种术语意义的混乱使得语言理论中的各个领域相互脱节，使得大多数语言学者只局限于自己的理论，有时还会错误理解其他的理论。

因此，在语言学研究中，准确使用和理解术语显得非常重要。因为术语中常常隐含着关于语言、语言本质及其他各方面的论述中可能存在的重大差别或相似之处。

20 世纪的语言学理论中，有人试图把各种东西联系在一起，还有些人

只在一个时间段内做某一件事。这不是原理问题,而是方法问题。

只要坚持运用所有理论中的任何一种,都有可能会取得成功,并产生深远的影响。这不仅影响语言教学理论,同时影响教材所依据的语言描写等。如果把语言看作一种状态,不同理论就会在不同的语法书和词典中对语言作不同的描写,从而决定教什么。如果把语言看作一种活动,也会产生不同的语言教学思想,决定如何教授语言。语言理论与语言教学是密不可分的。

## 参考书目

1. S·皮特·科德.《应用语言学导论》.上海外语教育出版社,1983
2. 陈菊先.语文教育学.华东师范大学出版社,1994
3. 朱绍禹.中学语文教学法.高等教育出版社,1988
4. 于根元主编.应用语言学概论.商务印书馆,2003
5. 中国大百科全书·语言文字卷.中国大百科全书出版社,1988
6. 吕叔湘.吕叔湘论语文教育.山东教育出版社,1987
7. W·F·麦基.语言教学分析.北京语言学院出版社,1990
8.《世界汉语教学》编辑部等.语言学习理论研究.北京语言学院出版社,1994
9. L·G·亚历山大.语言教学法十讲.科学技术文献出版社,1983

# 第二章　对外汉语教学

## 第一节　学 科 概 况

### 一、学科性质与名称

对外汉语教学是一门年轻的学科。它是语言学与应用语言学二级学科下的一个分支学科。其核心内容是将汉语作为第二语言教学。

由于是一门新兴的学科，因此在其迅速发展的同时，社会上、学术界乃至本学科内部对本学科的名称、性质、任务等基本问题尚有不同的看法或争论。所以，本学科的学科体系和理论需要进一步完善。本节先谈谈这方面的基本问题。

1. *学科名称*

(1) 语言教学中有关语言的几个基本概念

在讨论学科名称之前，我们首先要弄清楚几个与语言教学有关的语言的基本概念。

1) 第一语言和第二语言。第一语言和第二语言是按照人们获得语言的先后顺序来区分的两个概念，也是语言教学理论中用得最多的一对术语。第一语言是指一个人出生以后首先接触并获得的语言；第二语言则指人们在获得第一语言以后再学习和使用的另一种语言。当然有的人还会学习和使用第三、第四、第五种语言，但由于学习这些语言的规律与学习第二语言有很多共同之处，所以一般不再细分，而统称为第二语言（但也必须看到，第三、第四语言的学习和第二语言的学习者之间是存在差异的，比如学习目的语时所产生的正负迁移等）。同时也有这样的情况：有的幼儿同时习得两种或多种第一语言，达到同等熟练的程度，这就是所谓双语或多语现象。第一

语言、第二语言完全是从学习者学习语言的时间先后来区分的，这与一些多语言的国家由法律规定的第一语言或官方语言不是一回事。

2）母语和外语。母语和外语是按照国家的界限来区分的。母语指的是本国、本民族的语言；外语指的是外国的语言。一般情况下母语是人们的第一语言；但对一些移居国外的人的子女来说，他们出生以后首先接触并获得的语言有可能不是母语而是居住国的语言。因此，不能把第一语言和母语这两个概念完全等同起来。反之，第二语言也不一定就是外语。道理不言自明。

3）本族语和非本族语。本族语和非本族语是按照言语社团来区分的。通常是按民族的界限来区分的。本族语就是本民族的语言。因此这一术语与“母语”可以通用。非本族语是指本民族以外的语言，可能是外语，也可能是指本国其他民族的语言。比如，对于我国的各个民族来说，其他民族的语言是非本族语，但绝不能称作外语。从语言掌握的程度及运用情况来看，母语、本族语和第一语言通常是一个人的“主要语言”，但在有的情况下，也可能成了“次要语言”；外语、非本族语和第二语言一般是“次要语言”，但也可能成为“主要语言”。

4）目的语。目的语是指正在学习并希望掌握的语言。不论是外语或非本族语，甚至是非第一语言的母语，只要成为一个人学习并争取掌握的目标，都可以称为目的语。

以上概念是从不同的角度提出的，它们之间存在着意义上的交叉或模糊现象。比如，“第二语言”和“外语”这两个概念的区分就是一个比较复杂的问题。一般说来，第二语言是指母语（或第一语言）以外的、本国通用语或本国其他民族的语言，而外语则是指别的国家的语言。但是近年来出现一种用“第二语言”代替“外语”的趋势。广义地说，第二语言和外语的关系是包含与被包含的关系。

西方学者还从有无语言学习环境的角度对“第二语言”和“外语”这两个概念作了分工：凡是在该语言使用环境中学习的目的语称为第二语言；而不在使用环境中学习的目的语则称为外语。显然，从这个角度来区分第二语言和外语是有一定道理的，而且这种区分强调了有无语言环境对教学原则和教学方法等方面所产生的巨大影响，这也是很有意义的。

总之，“第二语言”有广义狭义之分。广义的“第二语言”是指任何一种在获得第一语言之后学习和使用的语言，包括外语；狭义的“第二语言”有两

种情况:一种是指第一语言以外的本国通用语或本国其他民族的语言,不包括外语;另一种是指在该语言使用的环境中学习的目的语,包括一部分外语。

(2) 学科的名称

一个学科的名称是该学科的内容和学科本质特点的反映。由于对某一学科本质的认识的不同,在学科名称上就会有不同的看法。目前在国内,本学科的正式名称是"对外汉语教学"。但也有一些不同的名称和看法。下面就作简单的介绍。

1) 对外汉语教学。这一名称基本上体现了这个学科的特点和内涵,在国内外也产生了广泛的影响,而且简洁上口,符合汉语习惯,因此从 1982 年提出一直使用至今。虽然有人认为这个名称似乎欠准确,但也没有人能够提出更能为大家所接受的表达法。中央文件、国家机构(如国家对外汉语教学领导小组)、学术团体(如中国对外汉语教学学会)等都正式采用这一名称。但是,20 多年的使用也发现这一名称有一定的局限性。即对外汉语教学这一名称只突出了主要的教学对象,但未能全面、准确地反映学科的性质——第二语言教学。不过由于该名称已被约定俗成地广泛使用,所以今后仍将是本学科用得最为广泛的名称。

2) 汉语教学。"对外汉语教学"本来是针对国内教外国人学汉语这一事业所起的名称,明显地带有以中国人的视角来指称这一学科的色彩。"对外"二字无法为国外从事汉语教学的同行所使用,因此,它只适用于中国。

海外从事汉语教学的人们根据各自的理解,给这门学科以不同的名称。有的把这门学科叫做"中文教学"(如美国),有的叫做"中国语教学"(如日本、韩国),也有的叫做"华文/语教学"(如东南亚国家,北京有"华文学院"、华语出版社、暨南大学有"华文学院")等。当国内外的学者在一起讨论学科或学术问题时,用"对外汉语教学"这一名称显然是不合适的。这种情况下一般使用"汉语教学"。正如设在北京的国际学术团体"世界汉语教学学会"及其会刊《世界汉语教学》,3 年一次已举行了 7 届的"国际汉语教学讨论会"等名称所表示的那样,由于语境清楚,一般不会与我国的汉语作为第一语言的教学(语文教学)相混淆,不会产生误解。这一名称在国际场合用得较多,已为各国学者所接受。国内学者一般不倾向于用"中文教学"、"中国语教学"、"华文/语教学"等海外学者所使用的名称,这与科学性及民族性有一定关系,因为"汉语"并不等于"中文"、"中国

语”、“华文/语”。

3）汉语作为第二语言教学。从学术上看，这一名称较为精确地指称了本学科的内涵和性质。它既能指在中国进行的对外国人的汉语教学，也能指世界各地的汉语教学，而且还能包括与之性质大体相同的对我国国内少数民族的汉语教学。也就是说，它能涵盖第一语言以外的所有汉语教学。现在这一名称在学术论著中使用得越来越多。但由于这一名称较长，不上口，再加上约定俗成的原因，它似乎不太可能取代“对外汉语教学”这个名称。不过今后在本学科的学术论著中它仍将继续出现。

**2. 学科的性质和任务**

(1) 学科的性质

对外汉语教学学科的核心就是对外汉语教学。其学科的性质也是由此而决定的。对外汉语教学是一种第二语言教学，即外语教学。

对外汉语教学是语言教学。语言教学的根本任务就是教语言，使学习者掌握汉语这一交际工具。它教授的是语言运用的技能，而不是语言学的知识和理论，因此它不是语言学教学。语言教学必然涉及并包括一定的文化内容，但对外汉语教学中的文化因素只能包含于语言教学之中，而不是文化包含语言教学，文化不能超越或凌驾于语言教学之上。

对外汉语教学是第二语言教学。这个性质将对外汉语教学与汉语作为母语的教学区分开来。众所周知，学校（无论是小学还是中学）里的母语教学都是在学生已经基本掌握了母语的听说甚至读写技能并具有相当的社会、文化背景知识的基础之上进行的。而第二语言教学则往往是从零起点开始，而且学习者缺乏社会、文化背景知识。这就使得对外汉语教学的目标、方法和教学侧重点与母语教学不完全相同。

对外汉语教学是汉语作为第二语言教学。对外汉语教学一方面受第二语言教学的普遍规律的制约，同时也受汉语本身的规律的制约，因此不同于英语、法语、俄语、日语等作为第二语言的教学。汉语在语音、词汇、语法特别是汉字方面有着许多自身的特点，这些特点往往成为教学的重点和难点，而这正是对外汉语教学区别于其他第二语言教学的地方。

对外汉语教学是针对外国人的第二语言教学。这与国内的少数民族的汉语教学有所不同。外国学生与我国的少数民族学生的差别在于前者学习汉语的社会、文化差异或冲突远远大于后者。因此对外汉语教学中的社会、文化因素必须予以足够的重视。

(2) 学科的任务

对外汉语教学的学科任务是研究汉语作为第二语言教学的原理、过程和方法,并以之指导教学实践。其具体内容包括:

1) 研究教学活动的主体,即研究学习者的生理、心理、学习策略和教师的基本素质;

2) 研究教学活动的客体,即研究作为第二语言教学的汉语本体的规律;

3) 研究教学活动的本身,即研究总体设计、教材编写、课堂教学和测试评估四大教学环节的理论和实践;

4) 将相关学科(语言学、心理学、教育学等)的理论应用于对外汉语教学研究并以自身的学科建设为相关学科的发展做出贡献。

## 二、学科的历史与现状

### 1. 我国对外汉语教学的历史发展

作为世界上最为古老的语言之一,汉语作为第二语言教学也有着悠久的历史。早在两千多年前的汉代就有一些国家派遣留学人员来我国学习汉语和文化。国力强盛的唐代是我国古代接受外国留学生最多的时期,其后的宋、元、明、清各代都有外国留学生前来学习。这些来我国学习汉语和文化的外国人主要来自我国周边的国家,最早到来的有印度、尼泊尔、巴基斯坦和中亚地区一些国家的僧侣,隋唐以后,包括如今日本、朝鲜、韩国、越南等国在内的周边国家不断有僧侣、使节、商人等来到中国学习汉语和文化;但清末直至新中国成立之前,由于战事不断、国力衰微,只有零星的外国人来中国学习汉语。

应该说,我国的对外汉语教学事业真正的发展是新中国成立以后的事。从 1950 年接受第一批外国留学生起至今,我国的对外汉语教学事业已经走过了半个多世纪的历程。这 50 多年的历史大致可以分为初创、巩固、恢复、蓬勃发展等 4 个发展阶段。

(1) 初创阶段(1950—1961 年)

1950 年我国正式接受东欧国家的交换留学生。在清华大学成立了东欧交换生中国语文专修班。这是我国第一个从事对外汉语教学的专门机构。1952 年暑期,随着全国高等院校的院系调整,清华大学东欧交换生中国语文专修班调整到北京大学,更名为北京大学外国留学生中国语文专修班。

20 世纪 50 年代还在广西桂林为越南和朝鲜留学生开办过中国语文专修班。

为了专门对大批非洲留学生进行汉语教学，1960 年 9 月北京外国语学院成立了非洲留学生办公室。1961 年，北京大学外国留学生中国语文专修班与北京外国语学院成立的非洲留学生办公室合并，更名为北京外国语学院外国留学生办公室。

这一阶段我国共接受 60 多个国家的留学生 3 000 多人。这些留学生在我国学习一至两年的汉语以后进入有关院校学习不同的专业，这时的汉语教学是一种预备教育。

此外，20 世纪 50 年代驻华外交人员的汉语教学也已开始，还开创了刊授和函授汉语教学。《中国建设》杂志于 1955 年开辟了“中文月课”栏目，厦门大学于 1956 年创办了华侨函授部，1957 年增设了中国语文进修班。

不少国家在向我国派遣留学生的同时，也要求我国派遣中文教师前去任教。50 年代我国曾向越南、匈牙利、保加利亚等国派遣汉语教师。国内培养对外汉语教师的工作也已开始。

初创阶段，在教学机构、师资队伍、预备教学体系等方面都是从无到有，为今后的发展奠定了一定的基础。

(2) 巩固阶段(1961—1966 年)

1962 年 6 月，经国务院批准，在北京外国语学院外国留学生办公室的基础上成立了外国留学生高等预备学校。该校的成立使我国的对外汉语教学有了一个稳定的基地，结束了教学机构和人员迁移不定的状况，是进入巩固阶段的标志。由于教学任务的不断扩大，该校的任务已不限于汉语预备教育，高教部于 1964 年将外国留学生高等预备学校更名为北京语言学院。迄今为止，北京语言大学(1996 年更名为北京语言文化大学，2002 年改为现名)仍然是我国惟一的一所以对外汉语教学与研究为主要任务的大学。

1965 年，越南政府向我国派遣了 2 000 名留学生，承担这批留学生汉语预备教育任务的单位，除了北京语言学院以外，还有北京大学、中国人民大学等 22 所院校。由于许多院校没有从事对外汉语教学的经验，高教部委托北京语言学院于 1965 年暑假为这 22 所院校的教师举办了一次对外汉语教学培训。这为后来的对外汉语教学事业的发展起到了很好的推动作用。为了加强各院校之间的交流，《外国留学生基础汉语教学通讯》也于 1965 年下

半年在北京语言学院创刊;这是我国第一份对外汉语教学的专业刊物。

从1962年到1965年,我国共接受外国留学生近4 000名,超过了前11年的总和。这一阶段对外汉语教学的其他教学形式也有所发展。中国国际广播电台于1962年开办了汉语教学节目。厦门大学华侨函授部于1962年扩充为海外函授部,教学规模迅速扩大。向国外派遣汉语教师的人数和对象国也有所增加。对外汉语教师的培养也得到进一步的落实。

由上述情况可以看出,到20世纪60年代中期,我国的对外汉语教学事业不仅得到了巩固而且具有了良好的发展趋势。但“文化大革命”的爆发使得高等院校全部停课,也不再招收外国留学生。中国国际广播电台的汉语教学节目,厦门大学的海外函授部的汉语教学也被迫中断。惟一的一份专业刊物《外国留学生基础汉语教学通讯》也不得不停刊。北京语言学院在停课几年之后被宣布撤销。从此全国的对外汉语教学中断了六七年。

(3) 恢复阶段(1972—1977年)

70年代初随着国际形势的发展和我国国际地位的提高,许多国家要求向我国派遣留学生。当时国内高校都已恢复招生,恢复接受外国留学生也势在必行。1972年6月北方交通大学首先接受了200名坦桑尼亚、赞比亚留学生。1972年10月周恩来总理亲自批示恢复北京语言学院。该院于1973年秋季开始招生。当年共接受42个国家的近400名留学生。随后,北京大学、复旦大学等一批院校也陆续成立了对外汉语教学机构,为在本校学习专业的外国留学生补习汉语。

这一阶段我国共接受外国留学生2 000多名,尚未恢复到“文革”前的规模。由于“文革”尚未结束,“极左”思潮仍然笼罩着全国,除了上课和编写急需的教材以外,教学理论研究工作还是难以正常开展。

(4) 蓬勃发展阶段(1978—1999年)

从1977年到20世纪末,是对外汉语教学的蓬勃发展阶段。这可以从以下几个方面来看。

1) 有了可靠的行政保障。由于我国与世界各国在经济、文化、教育、科技等方面的交往,随着我国改革开放的不断深入而日益密切频繁,人们认识到对外汉语教学的意义不仅仅是语言教学,它还有助于世界人民了解中国,促进人类的和平与进步事业。为了加强对对外汉语教学工作的领导和管理,协调各有关方面的力量,国务院于1987年7月批准成立了由7个部委参加的“国家对外汉语教学领导小组”。日常工作由其常设机构“国家对外

汉语教学领导小组办公室"（简称"国家汉办"）负责。

2）确立了对外汉语教学学科。1984 年当时的教育部长何东昌在一次报告中明确表示"对外汉语教学已发展成为一门新的学科"，国家教委在其后颁布的我国学科专业目录中列入了"对外汉语"这门新学科。1989 年国家教委在相关文件中明确指出"发展对外汉语教学事业是一项国家和民族的事业"。

1983 年成立"中国教育学会对外汉语教学研究会"，该会于 1988 年独立为"中国对外汉语教学学会"，至 2000 年已经举办了 7 届中国对外汉语教学讨论会；为了促进教学和研究，该会于 1996 年又先后成立了北京、华东、华南、华北、东北 5 个地区性的分会。1987 年成立了"世界汉语教学学会"，至 2002 年已经举办了 7 届国际汉语教学讨论会。

3）学术论文和出版刊物数量增加。1979 年我国第一个对外汉语教学的专业刊物《语言教学与研究》由北京语言学院的内部刊物改为正式出版的季刊。另两个对外汉语教学的重要期刊《汉语学习》和世界汉语教学学会会刊《世界汉语教学》分别于 1980 年和 1987 年创刊。这 3 个刊物是国内发表对外汉语教学论文的主要阵地，在国内外汉语教学界有着广泛的影响。相应地，本学科的论文无论是数量上还是质量上都是此前所不能比拟的。如果说此前的论文多为经验型、介绍型的，那么 20 世纪 80 年代以来，实证型、统计型定量分析的论文多了，跨学科、多角度研究的论文多了，研究汉语自身的教学和习得规律的论文多了。

4）办学规模迅速扩大。目前从事对外汉语教学的高等院校已经有 400 多所，到 2002 年底在校外国学生达 63 000 多人。而且许多学校都设立了独立建制的二级学院，这为教学和科研提供了切实的组织保障。

5）建立健全学科体系。20 世纪 70 年代末以前，我国的对外汉语教学基本上为汉语预备教育，属于非学历教育。以外国留学生为对象、以汉语作为第二语言教学为目标的四年制"汉语言"专业于 1975 年在北京语言学院试办，培养目标为汉语教师、翻译和汉语研究人才。其后南开大学、南京大学、复旦大学等院校也相继设立该专业，目前全国已有 30 多所院校设立了汉语言本科专业。

1986 年北京语言学院、北京大学开始在现代汉语专业中招收攻读对外汉语教学方向硕士学位的外国研究生，目前已有多所院校招收该方向的外国研究生。

1997年全国第一个带有对外汉语教学方向的语言学及应用语言学博士专业在北京语言文化大学设立。1999年,该校开始招收攻读对外汉语教学方向博士学位的外国学生,随后北京师范大学、中山大学、上海师范大学等院校也开始招收攻读对外汉语教学方向博士学位的外国学生。

因此可以说,我国的对外汉语教学经过长达50年的建设,终于有了从学士学位、硕士学位到博士学位的完整的学历教育体系。

这一时期教学大纲的研制、课程的设置和教材的编写,都由于办学规模的扩大和层次的提高而受到了国家汉办和各院校的重视。下一节将专门介绍这方面的情况。

**2. 我国对外汉语教学的现状**

新世纪以来,我国的对外汉语教学无论是在办学规模、办学层次、科学研究还是在对外合作和交流方面都有了更大的发展。我们从国家汉办的一系列新举措中就可以清楚地看到这些发展和变化。

1) 教材开发有了新的突破。2002年仅汉办组织编写出版的纸质和音像教材就分别达52种和19种之多。

2) 更加重视教学质量和评优工作。为了促进和优化全国的对外汉语教学,国家汉办于2002年首次举行全国对外汉语教学优秀教师评选活动。共有北京、上海、天津、南京和广州的11所院校的12名教师被评为全国对外汉语教学优秀教师。2003年又首次举行全国对外汉语教学优秀教材的评选工作。相信将有一批优秀的教材获奖并得到推广。

3) 重视基础和应用研究。国家汉办对本学科的科学研究更加重视,2002年启动的十五科研规划项目共有70多个项目立项。立项数目和资助经费都是前所未有的。同时还资助出版了10多种著作。

4) 国内教师外派和国外教师培训力度加大。随着国外汉语教学的升温,国外要求派遣汉语教师的高校越来越多,国家汉办或校际交流派遣出国的汉语教师也越来越多。在国内或在国外接受我国培训的外国汉语教师也越来越多。仅2002年在国内就为20多个国家培训了800多名汉语教师。与此同时,对外汉语教师资格考试已取消了教时限制;具有本科学历的人士均可报考。这是为对外汉语教学事业储备师资的新举措。

5) 建立对外汉语教学基地和汉语中心。自2002年起国家汉办准备用3年到4年的时间在全国建立十大对外汉语教学基地。这些基地在对外汉语教学领域应能起到示范带头作用,具有承担重大课题、编制教学大纲、编

写教材、远程教学、开发考试、培训教师、储备师资等功能且各具特色。经过2002年和2003年的两次评审，已经有北京语言大学、复旦大学、北京师范大学、北京大学、华东师范大学、南开大学、南京师范大学、中国人民大学等8所院校进入对外汉语教学基地的建设行列。与此同时，我国还准备在海外建立十大汉语中心，以发展海外的汉语教学事业，2002年已确定在纽约、温哥华、悉尼、汉城四个城市建设首批汉语中心。在国内则重点支持黑龙江大学、延边大学等10所院校开展面向周边国家的汉语教学工作。这方面的工作包括人员互访、情况调研、教师培训、中文竞赛、奖学金设立等。形成了政府与院校联合促进周边国家汉语教学发展的良好机制。

6）汉语水平考试(HSK)规模扩大，研发力度加强。HSK正在成为“世界上最具权威、影响最大的汉语水平考试”。2002年在国外参加HSK考试的人数近2万人，在国内参加的近4万人，历年参考人数累计已超过54万人。国内考点已达44个，国外也已在29个国家和地区设立了70个考点。此外，为了适应新形势的需要，在完善HSK主干考试的同时，国家汉办还组织相关院校研发HSK专项考试。目前已经确定的专项考试有：少儿HSK，商务HSK，文秘HSK和旅游HSK。其中少儿HSK已经进入预测阶段，2004年即将正式开考。

7）大力推广网络教学。国家汉办委托华东师范大学建立汉语远程学院。目前已经设立了网络电台，编辑了网络汉语杂志，研制了教学课件等。到2002年底已有128个国家和地区的4 600多名学生注册学习，汉语远程学院已经成为世界上最大的汉语远程教育机构。此外，国家汉办还积极承担了教育部中美网络语言教学合作项目的汉语教学部分。

综上所述，目前我国的对外汉语教学事业正伴随着祖国的发展而发展，形势一片大好。可以预见，随着我国经济的进一步腾飞、国际地位的进一步提高，我国的对外汉语教学事业必将更加辉煌。

## 第二节　教学大纲、课程设置及教材编写

### 一、教学大纲

教学大纲是设置课程、组织教学、编写教材和检查教学质量的基本依

据。各教学单位可根据不同的教学目标制订不同的教学大纲。如短期汉语教学大纲、预备汉语教学大纲、汉语言专业(本科)教学大纲等。教学大纲一般应包括：① 教学对象；② 培养目标；③ 教学内容；④ 教学安排；⑤ 教学原则；⑥ 课程设置；⑦ 教学要求；⑧ 附录。附录部分是各种教学参考资料和附表。如《高等学校外国留学生汉语言专业教学大纲》的附录即包括：词汇表、汉字表、语法项目表、功能项目表、课程介绍等。

由于在 20 世纪 80 年代以前，我国的对外汉语教学主要是非学历教育的预备教育和短期教学，因此制订全国性的汉语言专业的教学大纲未能及早地提上日程。直至 80 年代后期才出版了《汉语水平等级标准和等级大纲》(试行)(1988)。此后国家汉办、北京语言大学陆续组织研制了一些教学大纲，如《汉语水平词汇与汉字等级大纲》(1992)、《对外汉语教学语法大纲》(1995)、《中高级对外汉语教学等级大纲(词汇语法)》(1995)、《汉语水平等级标准与语法等级大纲》(1996)、《对外汉语教学初级阶段教学大纲》(1999)、《对外汉语教学高级阶段功能大纲》(1999)、《对外汉语教学初级阶段课程规范》(1999)、《对外汉语教学高级阶段课程规范》(1999)等。在此基础上，国家汉办组织全国力量于 2002 年制订出《高等学校外国留学生汉语言专业教学大纲》。这是一个全国各对外汉语教学单位和教师在教学活动中应该严格遵守认真执行的指导性文件。

## 二、课程设置

对外汉语教学的课程一般包括必修课、选修课和社会实践课等。根据不同的教学对象和不同的教学目标应设置不同的课程。课程设置就是在具体的教学目标和大纲的指导下，从学习者的特点和需要出发，根据专业对知识和能力结构的要求，最优化地选择教学内容、安排教学进度、形成配置合理的课程体系。比如，对短期进修教学、汉语预备教育和汉语言专业教学就应该设置不同的课程。下面对这三方面的课程设置分别作一点简单的介绍。

### 1. 进修教学的课程设置的原则

一般将学习期限在 1 周至 3 个月的汉语教学称为短期教学。来华进行短期学习的外国人的背景和学习目的各不相同。大致有：① 在校的汉语言专业大学生，他们一般利用寒暑假来华提高汉语；② 在校的其他专业的大学生或其他汉语爱好者，包括商业人士、教育工作者等，他们在假期中来

华是想了解中国、了解汉语。不管是哪一类人员,他们大都希望能运用汉语进行口头交际,而且大都带有旅游休假的目的,因此这类短期班的课程一般以听说为主兼顾读写,课程和进度也不宜安排过紧。同时应适当安排一些具有中国传统特色的课程,如民乐欣赏、传统体育、书画欣赏等,以满足学生了解中国的需要。

2. 汉语预备教育的课程设置的原则

汉语预备教育是为准备进入中国高等院校学习各类专业的外国留学生培训汉语。其目标是通过一学年的学习,使学生掌握必要的语言知识,具有一定的听、说、读、写能力和言语交际能力。了解一定的汉语言文化基础知识,能基本满足日常生活、学习和一般社交的需要。在学习完成后达到或超过进入文、理、工、医等专业学习所必需的标准。因为学习时间和学习目标明确,又因要在较短的时间内达到较高的目标,因此汉语预备教育的课程设置应具有快速强化的特点。快速强化具体体现在课时安排较多、课型较多,而且教学进度较快。下面是长期承担国家汉语预备教育的某校的课程设置。必修课程包括:汉语综合课、听力课、口语课、阅读课、正字正音课。选修课程包括:HSK 基础训练、HSK 强化训练、科技汉语、中国文化。辅助课程包括:传统体育、民乐欣赏、书法欣赏及实践等。另外,在第二学期(约在4 月中下旬)安排一次语言实践活动,时间约为一周。组织学生到文化旅游名城参观访问或到工厂、农村与工人、农民座谈、联欢,目的是检验前期教学效果,提高学生在现实的交际活动中的语言运用能力。要求学生努力使用汉语进行交际,实践结束后要求提交实践报告或游记一篇。

3. 汉语言专业教学的课程设置

将汉语作为第二语言教学的本科教学属于应用语言学,是外语教学专业,因此,专业培养目标和课程设置应与我国的其他外语专业相一致,而不是与国内的中文系相雷同。其培养目标为:本专业培养具有扎实的现代汉语基础和比较广泛的中国文化知识,能在教育、外事、商贸、旅游等部门从事教学、翻译、管理等工作的汉语专门人才。业务培养的具体要求是:本专业学生主要学习现代汉语、中国文学、中国历史、中国经济文化等方面的基础理论和基本知识,接受现代汉语听、说、读、写、译等方面的良好的熟练训练,掌握一定的科研方法,具有从事教学、翻译、研究、管理工作的业务水平及较好的素质和较强能力。

根据对学科性质的认识和定位,该专业的课程设置基本参照国家教育

部对外汉语专业课程设置的要求,主要注重语言能力的培养,同时开设文学文化课程。其中一、二年级主要是培养基本的语言技能,具体课程有:初级汉语综合课,初级口语,初级听力,正字课(选)和正音课(选)(一年级);中级汉语综合课,中级口语,中级听力,报刊语言基础,中级写作,书写技能(选),商贸汉语基础(选)(二年级)。三、四年级则语言、文学、文化并重,具体课程有:高级汉语综合课,高级口语,高级听力,高级写作,高级报刊语言(选),商贸汉语(选),中国概况,中国文化,古代汉语,现代文学史(三年级),现代汉语专题,论文写作,古代文学史,当代文学史,翻译与实践(选),汉语(作为第二语言)教学理论与实践(选),中文信息处理(选)(四年级)。另开设公共选修课:中国历史,中国哲学,现代汉语词汇,现代汉语语法,汉语中介语分析,书画鉴赏与实践及其他一些文化欣赏课。另外,在每年的春季(在4月中下旬)均安排一次为期一周的语言实践活动,以检验和提高学生在现实的交际活动中的语言运用能力。

## 三、教材编写

### 1. 教材的重要性和编写教材的依据

教材是教学计划和教学大纲的具体体现,是对外汉语教学培养目标、教学要求、教学内容和教学原则的具体体现,同时也是课堂教学和测试的重要依据。有无好的教材会直接影响教师水平的发挥和教学计划的完成。

教材除了在具体的教学活动中具有重要作用以外,对教学计划、教学大纲以及所依据的教学理论都有着重要的检验和反馈作用。通过教材的编写和使用可以引发编写和使用人员的理论思考,促进学科的发展。

教学计划和教学大纲是编写教材的直接依据这是不言而喻的。除此之外,编写教材还必须以当代的语言学、心理学和教育学理论为指导,以语言教学和习得理论为依据,这样才能保证教材的高质量。教材编写是科学性和艺术性很强的工作,任何轻视教材编写工作和马虎从事的态度都是要不得的。

### 2. 编写教材的目标

一部优秀的教材应该具备以下特点:

1) 针对性。教材要适合使用对象的特点。对于不同母语(文化背景)、不同年龄、不同目的、不同学习时限的学习者应有针对性的教材。

2) 科学性。教材要教授规范的语言,教学内容的编排要符合语言教学

规律，解释应准确规范，反映学科理论研究的新水平。

3）实用性。教材内容要从学习者的实际需要出发，尽可能贴近真实语言；要便于组织教学，便于体现交际化的教学原则。

4）趣味性。教材的内容和形式要生动有趣，尽可能使相对枯燥的语言学习过程变得轻松愉快。趣味性要和实用性、交际性结合起来，要防止低级趣味和文化冲突。

这些特点正是教材编写者应该努力追求的目标。在准备编写新的教材之前我们也应该思考，能否在这几个方面对现有的同类教材有所改进或超越。如果不具备这些条件，就不必匆忙编写，以免低层次的重复。低层次重复现象在现有的对外汉语教材中比较严重，应该引起重视。

3. **教材设计的类型**

对外汉语教材的类型可以从水平等级、教学类型和课程类型等不同的角度进行区分。这里主要从编排教学内容的角度介绍遵循不同的教学原则和编写体例所产生的不同类型。

（1）按遵循的教学原则区分

1）课文型。这种类型的教材是将课文作为教学内容的核心，通过课文学习词语和结构。课文大都为原文，具有真实的上下文和文化背景，有利于培养语言的运用能力，适合于中高级阶段。但由于不易系统、突出地进行语言要素和技能的训练，因此一般不宜据此编写初级教材。

2）结构型。此类教材是以语言的结构为纲，根据语法的难易程度安排教学内容和顺序。强调通过对结构（句型）的反复操练，养成语言习惯。缺点是对语言的交际性重视不够。

3）功能型。此类教材是以语言的功能为纲，根据功能项目的常用程度安排教学内容和顺序，较少考虑甚至不考虑结构的难易。这类教材的优点是强调培养语言的交际能力，但不可回避的缺点是对功能研究和切分还很不成熟，功能和结构的结合还没有得到较好的解决，往往忽视对系统的语言结构的掌握。

4）结构-功能型。这类教材是结构和功能相结合，以结构为主线，同时兼顾结构所表达的相应的功能。这样的安排主要是为了避免纯结构型和纯功能型的不足，兼收两者的优点。这类教材适合于初始阶段的语言教学，这也是改革开放以来我国对外汉语教材的主要编写思路。但这类教材也还没有很好地解决结构和功能结合的问题，结构往往成为支配因素，功能成为

点缀。

5）功能-结构型。与前一类型相反，此类教材在结构和功能的结合上由功能为主线，在一定的功能项目下安排结构教学。这种类型不容易兼顾语言结构的系统性，因此比较适合于已经学过基本的语言结构的中级水平的学习者。

结构-功能型和功能-结构型各有长短，适用于不同水平的教学对象。如何扬长避短、克服不足，是教材编写者需要进一步思考的问题。

6）话题型。话题型教材是以话题为纲安排教学内容，在话题中体现一定的功能和结构。这类教材选题比较灵活，能够体现一定的情景和文化，但功能和结构的系统性很难保证。中高级口语教材多采用这种类型的编排方式。

除了以上类型以外，还有以文化知识为纲编写的文化型教材。这类教材以介绍目的语文化为主的较容易编写，以双语或多语文化对比进行编写则具有相当的难度。如何将文化学习与语言学习很好地结合在一起，也还是正在探索的课题。

（2）按教材的体例分

1）综合型和分科型。综合型教材就是用一本（套）课本教授汉语。即在一本书里包括语音、词汇、语法等各种语言要素和听说读写等各种语言技能的学习。我国20世纪80年代以前的教材大多为这一类型。这类教材的优点是有利于综合性训练，缺点是各专项技能得不到专门的训练。而且一个班级只有一门课、一个教师，对教师的要求很高，也会影响学生的学习积极性。

分科型教材是按听说读写分技能编写的教材。80年代以后由于强调分技能训练而开设不同的专项技能课程，此类教材应运而生，出现了一批按技能分科的教材。如《初级汉语课本》、《现代汉语教程》等。这些教材有的以听说课本或读写课本为主干，其他课本与之配合；有的则是将语法和词汇作为共核，各科平行。这类教材的长处是既能使各项技能得到较充分的训练，又能使基本的语法和词汇在各科中得到复现和巩固。不足之处是往往容易产生各科教学分量不均、进度难以协调、相互牵制等问题。在一套教材中将听说读写密切配合的分科型教材目前多为初级教材，中高级阶段由于语言点的分散、词汇量的扩大和文化因素的增加，各门课程之间的相互协调较为困难，因此不宜采用整套的相互配合的分技能教材。

2）直线式和螺旋式。这是教材编排语法结构或话题内容的两种方式。

直线式就是将语法点按难易程度从第一课到最后一课依次排列(一课安排一至几个语法点),同一个语法点一般不重复讲解。对外汉语教材大都采用这一方法。这种方法的缺点是确定语法点的难易顺序往往较为主观,而且不符合语言学习需要循序渐进的规律。

螺旋式就是在课文的安排上将话题和语法结构根据交际需要由易到难构成几个循环。每个循环都逐步加大难度,表现为螺旋式上升。达到温故知新、稳步提高的教学目的。使用螺旋式的方法编排课文符合语言学习规律,应该是今后编写教材的主要方法。

## 第三节　教学流派和课堂教学

对外汉语教学的课堂教学技巧或方法,受到第二语言教学流派的直接影响。因此本节首先介绍第二语言教学的主要流派,然后讨论各种课型和各语言要素的教学技巧。

### 一、第二语言教学流派简介

现代外语教学只有 100 多年的历史。100 多年来,先后出现了数十种教学法流派。其中影响较大的有强调自觉掌握的认知派,如语法翻译法、自觉对比法和认知法等;强调习惯养成的经验派,如直接法、情景法、听说法和视听法等;强调情感因素的人本派,如团体语言学习法、默教法、暗示法等;强调交际运用的功能派,如交际法。下面概要予以介绍。

**1. 语法翻译法**

语法翻译法(grammar-translation method)是最古老的外语教学法,在欧洲用来教授古希腊语、拉丁语已有千百年的历史。它是以系统的语法知识的教学为纲,依靠母语,通过翻译手段,主要培养外语读写能力的教学法。德国语言学家奥伦多夫等于 18、19 世纪之交对其进行了理论上的总结和阐述,使之成为一种科学的外语教学法。这种教学法的语言学基础是历史比较语言学,心理学基础是官能心理学。其主要特点为:

1) 以书面语教学为主要目标,不重视口语和听力教学;

2) 以系统的语法知识为主要的教学内容,采用演绎法详细地分析语法

规则并通过翻译练习加以巩固；

3）强调学习规范的书面语，注重名著和原文；

4）用对译的生词表进行词汇教学，句子是讲授和练习的基本单位；

5）用母语教学，翻译是主要的教学、练习和评估手段。

语法翻译法是外语教学法历史上第一个完整的教学法体系。它强调对语法规则的掌握，注重学生智力的发展，能较好地培养阅读和翻译能力。主要不足是忽视口语教学，过分依赖母语和翻译，过分强调语法教学。

2. 直接法

直接法(direct method)19 世纪末 20 世纪初产生于西欧，是与语法翻译法相对立的教学法。这种方法是直接用外语教授外语，而不用母语和翻译，也不注重语法形式；其教学目标是口语而不是规范的书面语。其主要特点为：

1）教学中排除母语和翻译，采用各种直接的手段用外语学习外语；

2）以口语教学为基础，先听说后读写；

3）以句子为教学的基本单位，整句学整句用，在句子中学习单词和语音；

4）以直接感知、模仿为主要教学手段，到一定阶段对语法进行总结和归纳，而不是先教语法；

5）以当代通用的活的语言为教学对象，而不是教授名著中典雅但已过时的语言。

直接法的这些主张丰富了人们对外语教学规律的认识，对后来出现的一些教学法，如听说法、视听法、自觉实践法、功能法都产生了很大的影响。

但直接法也存在一些缺陷：

1）对成人学习外语的规律认识不足，将儿童习得母语和成人学习外语混为一谈；

2）强调口语教学，不够重视读写能力的培养；

3）过分强调模仿，偏重经验，不够重视人的积极学习的主动性。

3. 情景法

情景法(situational language teaching)是在直接法受到越来越多的批评的情况下于 20 世纪二三十年代在英国产生的。它以培养口语能力为基础，强调在自然情景中教授语言。其主要特点是：

1）教学目的是对各种语言技能的实际掌握，语言技能又是通过对语言

结构的掌握而获得的；

2）先教口语再教书面语，学生有了一定基础以后再教阅读和写作；

3）课堂用语为目的语；

4）新的语言点在情景中介绍、操练；

5）运用词汇选择程序，保证基本词汇的教学；

6）按照先易后难的原则对语法项目进行分级排序。

情景法继承了直接法的长处，同时又在对课程内容的系统选择、安排和呈现原则、控制词汇和语法，特别是强调运用情景方面，较直接法有了很大的发展。

4. 听说法

听说法(audiolingual method)产生于20世纪40年代的美国。该方法强调通过反复的句型操练培养听和说的能力。其特点为：

1）听说领先，读写跟上；先教口语，后教书面语；

2）用模仿、重复和记忆的方式学习，以形成习惯；

3）以句型为教学中心，通过句型操练掌握目的语；

4）通过母语和目的语的对比确定教学难点；

5）严格及时地纠正学生的错误，避免养成错误的习惯；

6）大量使用录音、电影、电视等现代化教学手段进行强化刺激。

听说法继承了直接法的主要教学原则，同时又有所发展，特别是强调在对目的语和学生母语进行对比的基础上编写教材与进行课堂教学。听说法至今仍在美国外语教学中占据重要的位置，并在许多国家的外语教学中继续发挥作用。代表性教材是《英语900句》。但该法也存在一些明显的不足：

1）过分强调结构的机械训练，而这些结构往往脱离语境和内容，枯燥无味，引不起学生兴趣，难以培养语言的交际能力；

2）注重听说，忽略读写可能使学习者的阅读、写作能力较差，知识面较窄，缺乏后劲；

3）目的语和学生母语的对比，并不能预示学生的全部难点和错误。

5. 视听法

视听法(audio-visual method)20世纪50年代产生于法国。它是强调在一定的情景中听觉感知和视觉感知相结合的教学方法。该方法来源于直接法和听说法，继承和发扬了两者的长处，克服了听说法脱离语境机械操练

的不足，创造性地提出了视觉听觉相结合的整体结构感知的新教学法体系。其代表性教材是《新概念英语》。视听法的主要特点是：

1）利用电化教学手段使视觉感知和听觉感知相结合；

2）创造接近于真实的情景使语言和情景紧密结合；

3）通过图像和完整的对话使语言要素在对话中被整体感知，教学顺序是先整体后部分：成段对话—句子—单词—语素；

4）先教听说后教读写。

视听法也存在一些不足：

1）重视口语而忽视阅读训练；

2）过分强调整体感知和综合训练而忽视语言结构的分解和单项训练；

3）只注重语言形式，未能着重培养交际能力——这是所有结构法的共同弱点。

6. 认知法

认知法(cognitive approach)20 世纪 60 年代产生于美国。其主张和听说法相对立，而且对强调语法教学和发展智力的语法翻译法予以肯定，因而被称为“现代语法翻译法”。其语言学基础是乔姆斯基的转换生成语法理论，认为语言是一种受规则支配的体系而不是习惯体系，人类学习语言是一种有意识的创造性的运用过程，而决不能仅仅依靠模仿记忆。其心理学基础是认知心理学，强调学习不是简单的刺激-反应和模仿重复，而是一个感知、记忆、思维、想像的过程，是大脑抽象思维活动的结果。认知法有如下特点：

1）强调发挥学习者的智力作用，通过观察、记忆、思维、想像等活动，获得正确运用语言的能力，反对动物型的刺激-反应式的学习；

2）强调在理解掌握的基础上进行有意义的练习，提倡演绎式的教学，反对机械模仿；

3）以学生为中心，注重培养学生的学习信心、学习兴趣和自学能力，充分调动学生的学习积极性和主动性；

4）主张听说读写同步推进，从初始阶段就进行全面训练，口语和书面语并重；

5）适当地使用学生的母语，通过对比确定教学难点和重点，可用母语解释比较抽象的语言现象，但反对滥用母语；

6）正确对待学生的错误，反对有错必纠，防止因过多的纠错使学生怕

出错而不敢运用甚至丧失学习信心。

认知法的最大贡献是更多地从心理学的角度研究外语教学。另外,它从不同的甚至对立的教学法流派中吸取有益的成分并克服其片面性,也是其影响较大的一个原因。

除了以上介绍的教学法以外,值得一提的还有阅读法、自觉对比法和自觉实践法。前者产生于20世纪初,是一种强调通过直接阅读来培养阅读能力的教学法;后两者分别产生于20世纪30年代至50年代和60年代的苏联。自觉对比法是一种主张通过母语和目的语的翻译和结构对比,自觉掌握目的语的教学法,是一种与直接法相对立的教学法。自觉实践法则摈弃了自觉对比法过于强调翻译对比和理论知识的讲解,忽视操练和交际性练习的做法,继承了直接法通过大量的语言实践培养听说读写技能的特点和语法翻译法强调自觉掌握语言规则的主张,成为一种较具综合性也较成熟的教学法。以上介绍的是认知派和经验派教学法,下面是人本派和功能派教学法。

7. 团体语言学习法

团体语言学习法(community language learning)20世纪60年代初产生于美国,是一种采用小组集体讨论的形式、教师和学生之间类似医生和病人的关系并把学习过程看成是咨询过程的外语教学方法。其突出的特点是以学生为中心,教服从于学;另一个突出的特点是依靠学生的母语,逐渐由学生的母语过渡到目的语。这一教学法也存在明显的不足:其一是教师的指导作用发挥不够,教学内容和形式一切都由学生自己决定,往往成了学生牵着教师的鼻子走;其二是这种教学方式对教师的要求更高,教师必须具有较高的业务水平和丰富的教学经验,要能够因势利导、随机应变,这是一般的教师难以做到的。

8. 默教法

默教法(the silent way)是60年代初在美国创立的教学法。它要求教师在课堂上尽量少说话,尽量鼓励学生多参与语言活动,从而使学生更有效地掌握运用外语的能力。其主要特点是:

1) 以学生为主体,教服从于学;

2) 教师的沉默是其最大的特点,随着学生水平的提高,教师说得越来越少;

3) 教师一般不改正学生的错误而让学生自己改正错误;

4) 口语领先,随后培养读写能力;

5）强调学生之间的合作而不是竞争，提倡互相帮助，共同解决问题；

6）把词汇看作语言学习的核心，将词按功能进行分类，按照一定的顺序学习。

默教法的最大特点也是它的不足之处，过分强调教师少说话，使得教师的指导作用不能正常发挥，这不仅影响语言知识的传授，也影响教师的示范作用。

9. 暗示法

暗示法(suggestopedia)20 世纪 60 年代中期创立于保加利亚。该方法强调通过暗示激发人的潜力，使有意识和无意识的活动相结合，让学生在放松而又注意力高度集中的心理状态下有效地学习。暗示法的教学特点有：

1）布置舒适幽雅的环境，创造轻松和谐的气氛，以此消除学生的紧张心理，激发其记忆力和想像力；

2）充分发挥无意识活动的作用，做到有意识和无意识活动的协调发展，以开发学生的学习潜力；

3）建立高度的自信心，鼓励学生大胆运用目的语，尽量少纠错，没有强制性的课外作业，更不给学生考试的压力，以免造成紧张的心理；

4）在消除学生的心理障碍，充分发挥其潜力的情况下，扩大语言材料的输入量。

暗示法强调在外语教学中运用心理学方面颇具特色，但对语言知识的系统传授及其对语言交际的指导作用则重视不够，而且语言学习中是不是应该完全排除紧张和焦虑情绪也还值得探讨。

10. 全身反应法

全身反应法(total physical response)这种教学法 20 世纪 60 年代初期在美国产生，70 年代盛行。它强调语言学习行为的协调，通过身体动作教授外语，主要用于美国移民儿童的英语教育。全身反应法以儿童习得母语的理论为依据，将听力理解的重要性放在第一位。通过动作指令学习语言，是这一教学法的一大特色，但动作表达语义受到一定的限制，因此这种方法用于外语教学的初始阶段较合适，而在中高级阶段不宜单独运用。

11. 自然法

自然法(natural approach)是在 20 世纪 70 年代后期产生于美国的新的教学法。该法提出在非自然的条件(即课堂教学)中最大限度地扩大对学生的语言输入，首先集中培养理解能力并强调通过习得掌握外语。其特点是：

1）教学目的是培养学生的交际能力，课堂活动全部用于交际；

2）强调理解的重要性，认为语言学习应从理解开始，加强输入（听和读）的活动，推迟输出（口语表达）；

3）强调自然习得，在课堂上尽可能提供语言习得的机会；

4）教师一般不纠错，强调学生自我纠正；

5）创造轻松愉快的学习气氛，根据学生的需要和兴趣确定教学内容。

自然法在很多问题上，特别是在非自然环境的课堂教学中如何进行自然习得给人们很多启示。但该法也存在一些争议，主要是过分强调习得的作用，而其过分依赖的“自然顺序”假说尚未完全证实。

12. 交际法

交际法（communicative approach）早期被称为“功能法”、“意念-功能法”，是以语言功能和意念项目为纲、培养能够在特定的社会语境中运用语言进行交际的一种教学法。

交际法20世纪70年代初期产生于西欧，中心是英国。其影响最大的教材是《跟我学》。交际法一反其他教学法以结构为纲的传统做法，明确提出以功能意念为纲。该方法并不否认语言知识的教学，也并不排斥适当的词汇和句型操练；只是强调语言形式是实现语言功能的手段，结构形式的教学要服从于语言功能。交际法的特点是：

1）明确提出外语教学的目标是培养创造性地运用语言进行交际的能力，不仅要求语言运用正确，而且要求得体；

2）以功能和意念为纲，根据学生的实际需要，选取真实自然的语言材料；

3）教学过程交际化，创造接近真实交际的语言环境，教学方法多样化，而且多为实用的交际活动；

4）以话语为基本的教学单位，语音、词汇和语法主要通过话语情景综合教学；

5）单项技能训练与综合性技能训练相结合，不排斥单项技能的训练，但要以综合性训练为主；

6）强调表达内容，不苛求于语言形式，对学生在学习过程中出现的错误要有一定的容忍度，不影响交际的一般性错误能不纠正就不纠正；

7）按话题螺旋式地安排语言材料，将功能和结构统一在话题中，每一话题有几次循环，每次循环增加一定的难度。

交际法博采众长，从当代语言学和心理学研究的最新成果中获取营养，同时又吸收了之前的其他教学法的长处并克服其不足；因此，交际法成为迄今为止影响最大也最有生命力的教学法。但因为交际法提出的时间并不长，也还存在不少需要解决的难题：

1）交际法所依据的语言功能和意念的基础研究还很不够，语言功能和意念是很宽泛的概念，目前还没有像词汇、语法大纲那样比较成熟的功能大纲和意念大纲，对它们的确定、分类和排序还缺乏科学的依据；

2）语言功能和结构在教学中如何协调配合是长期以来未能很好解决的问题，交际法以功能意念为纲，这往往难以照顾到语言结构的系统性，难以很好地安排语言结构的循序渐进；

3）难以真正做到课堂教学的交际化，课堂教学的交际化对教师的要求过高，一般教师尤其是非本族语的教师很难做到，而且如何通过课堂中的交际活动培养交际能力，如何在这样的教学活动不忽视语言基本要素的掌握都还需要探讨；

4）交际法的变革，给教师培训、教材编写和测试评估等等都提出了新课题，缺乏相应的教师、教材和测试评估体系，交际法的实施就缺乏必要的保障。

我们应该看到各种教学流派的产生都有其必然性与合理性；它们各有长处，也都各有不足；相互之间不仅存在着对立和排斥，同时也有继承和发展；新老教学法并不是互相替代，而是长期共存；而且不同流派在保持自身特色的同时，出现了综合化的趋势。

## 二、课堂教学

### 1. 课堂教学的重要性

课堂教学是对外汉语教学的基本组织形式，是实施教学计划、贯彻教学原则、运用教学方法、完成课程教学、实现教学目标的依靠。对这些方面的检验和评估以及对教师素质和能力的评价都离不开课堂教学。因此可以说，课堂教学是教学活动四大环节中的中心环节。任何教学单位和教师都应该充分重视课堂教学。

### 2. 课堂教学的基本环节

这里介绍不同的课型所应具备的五个基本的教学环节：组织教学、复习检查、讲解新课、巩固新课、布置作业。

1）组织教学。本环节的目的是使学生情绪稳定，注意力集中，做好上课的思想准备。这一环节较短，占三五分钟。开始上课时，师生互相问候，可随便谈论共同感兴趣的话题以活跃气氛；然后考勤检查出席情况，询问缺席学生的情况以示关心；接着介绍本节课的教学目标和内容，使学生对其有个大致的了解。

2）复习检查。本环节的目的是了解学生对已学内容的掌握情况，获得反馈信息，督促学生课外的复习，为下一个环节打好基础。不同的课型可有不同的复习检查方式，具体的做法很多。这里仅介绍一些常用的方法。对初级班的语音和词汇，可采用听写的方式，听写时可让部分学生在黑板上写，以便及时批改；中高级班的词汇，除了听写以外更常用的方法是教师释义让学生说出生词；语法或功能项目可采用造句或会话的方式；对课文内容可采用复述、演讲的方式；如有课外批改的作业则应对有共性的问题加以讲解。

复习检查在整个课时中所占的时间并不多，一般在10～15分钟。所以复习检查的内容应该是上节/次课所学习的重点而不是全部内容。另外，在复习旧课导入新课时如果能够做到新旧内容的自然衔接，则可收到温故知新的良好效果。

3）讲解新课。这是课堂教学的主要环节，是学生接触获得新知识增强技能的环节。新课涉及语音、生词、语法、修辞和课文、练习等等内容。在这一环节中特别要注意教师与学生的互动，即要充分调动学生参与的积极性；不允许教师不顾学生感受地“一言堂”。可以将各语言要素分割成小单元，讲练结合；尽量避免教师的长篇大论。比如处理生词，可以让一个学生先读，然后教师纠正；再一个学生读，教师再纠正；接着教师带读；然后再解释重点词语。这样就比只是教师一个人讲解要好。

4）总结新课。在一次课结束之前应该用10～15分钟的时间对本次课所学习的新的语言点加以归纳总结。总结可以有两种方式：一种是由教师进行条理分明的总结，这不只是回顾本次课学了点什么，更重要的是要归纳出语言点的特点，比如概括成结构或句式，指出学生容易出错的方面等；一种是让学生回忆总结，这时需要教师的引导、提醒和概括。

5）布置作业。布置课外作业是教学环节的最后一环，但不是可有可无的。它是巩固课堂教学内容、督促学生及时复习和预习、保证教学进度和质量的重要手段。在下课前应留下几分钟的时间布置课外作业。课外作业的

内容应该与当天讲解的内容、下一次的课堂复习及新课的内容相联系，使各个教学环节环环相扣，形成有机的整体。

3. **课堂教学的技巧**

教学目标和教学原则要通过具体的教学方法和技巧来实现。同样的教学内容运用同样的教学原则和方法，但教学效果不一定相同，这其中教学技巧起着很大的作用。

（1）备课和教案

备课是课堂教学的基础。没有课前的充分准备是不可能有良好的课堂教学效果的。教师在接受一门课程的教学任务时，首先应该明确教学大纲的要求和课程的教学目标，了解教材特点。据此制定学期教学计划，包括教学重点、学时分配和每周进度。在制定好教学计划以后，即可备课并写出教案。教案可按教材的一课或一个单元来写，同时要大致确定各部分所需的学时——实际上课所用的时间可能与教学计划、教案上所设想的时间有些出入，这可以适当调整，但出入不应太大。有三个方面的问题是备课时必须注意的：

1）分析教材。教师应细致分析教材，掌握教材以及各课的教学目的；要了解平行课本与本教材的关系——各自的侧重点是什么。据此确定教学的重点和难点。

2）分析教学对象。除了了解掌握教材情况以外，要备好课还应对学生的状况有所了解。这包括学生的现有水平、学习动机和态度、学习和交际策略等等。教师应经常用口头询问的方式——必要时也可以采取问卷调查的方式——了解学生情况，征询学生对教学重点、教学进度及教学方法等方面的意见。

3）确定教学方法。在了解教材和学生的情况之后，才能选用具体的教学方法。展开教学的方式——是从课文开始还是从句型结构开始，是将生词、课文、语法分开讲还是在课文中讲解生词和语法，是用归纳法还是用演绎法；安排机械式练习和交际性练习的时间和方式；以及如何提问，如何利用教具和电化设备等等。这些都应该在教案中写明。

教案内容一般应包括课程名称、教学内容、目的和要求、重点和难点、学时安排、所需教具、教学环节、板书设计、教学后记等。其中教学后记应该予以特别的重视，它可以记录下教案在实施过程中的得失，为改进教学提供参考。教案的详略可根据教师的经验和每课的特点决定。一般而言，新教师

的教案应该尽可能地详细一点。

(2) 课堂的组织

要成功地组织课堂教学,必须使课堂气氛活泼有序。要做到这一点,必须充分发挥教师的主导作用和学生的积极性。

要创造轻松活泼的课堂气氛,就必须消除学生可能出现的畏难、压抑、焦虑、厌烦等不良情绪。因此,教师的亲切热情、幽默风趣的教态起着很大的作用。但活泼并不是嘻嘻哈哈,轻松也不是松松垮垮。应该处理好课堂气氛的活泼与严肃、轻松与紧张之间的辩证关系。

要自始至终驾驭课堂,吸引学生的注意力。除了教学内容生动有趣实用、教学环节张弛有序以外,教师自己应该精神饱满,全神贯注,而不能精神涣散,漫不经心;教师的语言和表情动作要有感染力,要面向全体学生交流,而不能表现出个人的好恶。

要能够稳定课堂的正常的教学秩序。对交头接耳、思想开小差、东张西望、疲倦打盹等不认真上课的行为不能听之任之,要给以适当的提醒。防止这些行为的方法有很多,下面略举几种。就某一个学生陈述的内容提问——这可以用于一个学生说话,别的学生不注意听的情况;让几个学生互相衔接朗读一段文章——可以突然让正在朗读的学生停下来,让思想不集中的学生接着读;也可以在讲解或练习当中突然提问。

(3) 教师的语言和提问

对外汉语教师的课堂教学语言一般而言有三大特点:第一,语速较慢,停顿较多;第二,使用的词汇、句子较简单;第三,冗余信息较多,即解释、重复较多。但教师的语言有示范和引导作用,要推动学生向目的语靠拢,因此必须是规范的。不能为了降低学生理解的难度而使自己的语言在语音、语法等方面出现不规范的现象。因此,教师的教学语言应该注意以下几点:

1) 不论语音、词汇或语法,都应该准确规范;

2) 除了语音操练或新词语教学以外,无论课堂教学还是课外交流,语速都应该正常或接近正常,避免夸张;

3) 选词用句要浅显易懂,随着学生程度的提高逐步增加语法的难度和词汇的多样化;

4) 要有意识地重现已经学过的句式和词语,避免使用尚未接触到的词汇和语法;

5) 尽量用汉语和学生交流,避免使用或夹杂学生母语或其他媒介语;

6）应克服口头禅或过多的重复、啰嗦等其他不良的语言习惯。

课堂上的提问是教师组织教学的重要手段，也是教学技巧的体现。教师一般应该面向全体学生提问，让所有的学生都有回答的机会。对水平较高、发言积极的学生和能力较差、不愿举手的学生要有所调控，尽量做到机会均等。一种有效的方法就是教师提出问题请大家思考，然后指定一个学生（往往是不太主动、机会较少的）回答，待他回答以后，其他同学可以补充。

从教师提问到学生回答，应该用多长的等待时间？这是一个值得探讨的问题。有的教师提出问题后，让学生马上回答出来，如果学生不能立刻回答就随即换人或自己立刻把答案说出来，这样既起不到提问的作用又可能损伤学生的积极性和自尊心。提问以后应该给学生一点思考的时间，如果学生表示答不出来，可以考虑换人。提问以后教师也不能一味地消极等待，使得其他学生无所事事，课堂气氛沉闷。如果被提问的学生还在思索，教师可以关切的询问："想好了吗？可以回答了吗？"以缓解沉闷的气氛。

（4）板书

板书是保证教学质量的重要教学手段，它的作用是使教学纲目清楚、重点突出，帮助学生理解和记忆。因此，教师对板书这一教学手段应予以充分的重视，精心设计的板书不仅能够吸引学生、促进教学，而且也能反映教师认真、严谨、艺术的教学风格。板书的基本原则是：

1）板书应精心设计。可以将黑板分成两大块，左边书写教案中准备好的内容，如本课的重点词语和句型；右边用于教学中临时性的板书。板书，尤其是课前准备好的板书应该条理分明，切忌杂乱无章。

2）汉字书写正确、规范、美观。教师的板书有示范作用，因此，笔画、笔顺一定要符合规范，不能将个人不规范的书写习惯带到课堂上。有的教材上的少数笔顺也不合国家规范，教师应该能够辨别，不应将错就错。汉字书写也要力求美观，至少也要端正。

3）板书应该鲜明、醒目，重点部分可以用彩色粉笔标出。比如汉字的重点笔画、句型特别需要注意的地方等。

**4. 课堂教学的评估**

课堂教学评估，对于教师和学生、对于教学同行和管理者都很重要。教学同行和管理者可以通过教学观摩，对课堂教学展开研讨，优化教学；教师

可以从教学评估中获得反馈，总结得失改进教学；学生在教学评估中反映出的意见和要求对完善教学也有一定的作用，不应忽视。

课堂教学评估应该从教学目的、教学内容、教学环节、教学技巧等方面对课堂教学的全部活动进行全面评估。大致的内容如下：

1）课堂教学的目的是否明确。比如，口语课是否着重训练口语，阅读课是否着重训练阅读能力等；

2）是否做到教学内容重点突出、难点分散；

3）教学环节是否安排得紧凑、自然、合理。比如，复习与新课是否有机地衔接；

4）时间的分配是否得当。比如，讲练时间的分配、生词处理与课文串讲时间的分配等；

5）教学技巧是否娴熟自如。包括驾驭课堂的能力、教态和课堂气氛、提问的方式、板书的质量、辅助教学手段的运用等。

6）能否完成教学大纲所规定的教学任务，教学效果如何，即教学进度与大纲要求是否一致，学生能否学以致用。

## 第四节　第二语言习得

### 一、对比分析理论

#### 1. 理论对比分析和应用对比分析

有关第二语言习得的理论和模式，西方学者已经提出了许多种。我们这里主要介绍目前国内对外汉语教学界比较熟悉并应用较多的对比分析理论和中介语理论。

对比分析也称对比语言学，就是将两种或两种以上的语言进行共时对比，来确定其相同点和相异点。

对比分析又分为理论对比分析和应用对比分析。

理论对比分析的任务是：通过语言的对比，对语言的相同点和不同点进行全面的说明和解释，并且提供语言对比的理论模式，这种对比要求详尽、明确、严格。

应用对比分析是应用语言学的一个分支。它运用理论对比分析的研究

成果,提出语言对比的框架,为语言教学提供必要的信息。比如:选择语言中可能给语言教学带来困难的语言项目进行对比;向语言教师、教材设计者提供充分的信息和数据,使他们更有针对性、更有效地制订大纲,设计课程,编写教材,改进教学。

应用对比分析作为一个专门学科,在国外已有近半个世纪的历史,而在我国却起步较晚。较早进行这方面研究的有吕叔湘、王还等先生。

对比分析虽然重在语言形式(包括语音、语法结构、词汇等),但它对语言教学具有不可忽视的指导作用,是每一个语言教师必须具备的基本功。

对比分析是一门既古老又年轻的学科。自从有了语言之间的接触,就有了语言的对比。但它作为一门专门的学科而应用到语言教学中来还是从拉多(R. Lado)开始的。拉多的《跨文化的语言学》(1957)对外语教学影响很大。

对比分析之所以成为可能是基于:各种语言之间有着很多相同和不相同的地方,这些异同点是可以描述的,它们也是人们学习第二语言时产生迁移的根源。

对比分析将迁移分为正迁移、负迁移和零迁移。两种语言相同的,可能发生正迁移;不完全相同但有关系的,发生负迁移;完全不同没有关系的,就不产生迁移即零迁移。学习第二语言就是要利用正迁移,克服负迁移。

负迁移也叫干扰。它又可分为两种,一种是阻碍性干扰,即目的语中有而母语中无的东西会干扰新东西的学习;一种是介入性干扰,即目的语中没有而母语中有的东西对学习第二语言的干扰。

**2. 对比分析的步骤**

对比分析是双向的,而对外汉语教学中的对比分析的目标是单向的——通过对学生的母语和目的语的对比分析,找出学生学习汉语的难点和重点,预计学生可能产生的困难和问题。难点和重点是根据学生对象确定的——因此,亟需编写针对不同母语学习的专门性教材,而不是通用性教材——当然,专门性教材和通用性教材都是建立在不同的理论假设的基础之上的,前者认为不同的母语产生不同的负迁移,后者认为不同的母语背景的学生的偏误存在共性。问题是母语的负迁移和学习共性各自所占的比例究竟有多大?

确定难点和重点并不意味着对比分析是分散的、零碎的。应用对比

分析也是系统的、全面的，目的语中的难点和重点是在对两种语言的语音、词汇和语法进行系统对比的基础之上确定的——即使我们进行个别的或局部的对比分析也应该有系统的全局的观点，否则就容易顾此失彼，捉襟见肘，所以，我们首先得有汉语本体研究的基本功和扎实的外语功底。

经典性的对比分析一般按以下步骤进行：

1）描写。对学生的母语和目的语进行准确、清晰的描写。这种描写以一定的语法体系为依据；

2）选择。选择一定的语言项目、规则或结构进行对比；

3）对比。找出两种语言关系中的特殊点。这要依靠参照点的有效性——即具有可比性；

4）预测。形成预测的错误和难点。这种预测可以通过难度等级模式和心理学、语言学理论的应用而得到。

为了使对比分析的预测能够形式化，排除主观因素的干扰，应用语言学家建立了各种难度等级模式，对学习中的难度进行了分类。现在经常介绍到的是布拉图(或译为普拉克特，Cliford Prator, 1967)提出的较为简明的难度等级模式。他将难度等级分为6个等级：

1）零级：迁移(transfer)。两中语言里相同的成分，包括语音、词汇和语法，产生正迁移。没有区别，学习没有困难。如"主谓宾"结构等。

2）一级：合并(coalescence)。母语中的两项，到目的语中合并成一项，学生可以忽略原来的差别，习惯新的项目。例如英语的第三人称单数和汉语中的第三人称单数就是二对一的关系。

3）二级：差异不足(under differentiation)。母语中有而目的语中没有的语言项目，学生必须避免使用，以防止介入性干扰。比如，汉字书写中的母语干扰。

4）三级：再解释(reinterpretation)。母语中的某个语言项目在目的语中虽有相应的项目，但在项目的形式、分布和功能方面又有差异，学习者必须把它作为新项目重新习得。

5）四级：超差异(over differentiation)。目的语中的某个语言项目，在母语中没有相应的项目，学习者在习得这些项目时会产生阻碍性干扰。如汉语中的汉字、声调、翘舌音等多是其他许多语言所没有的。

6）五级：分裂(split)。即母语中的一个语言项目在目的语中分裂成两

个以上的项目。这一级正好跟一级(合并)相反,需要学生细加区别。

难度等级模式的建立,有助于预测的形式化,减少预测中的主观色彩。由于该模式具有很强的逻辑推理性,应该说,这个模式完全可以通过推理设计出来,而不一定需要多少实例的归纳。实验在多大程度上支持这样的模式,还需要进一步研究。汉语教学的难点是否与该模式一致也还有待于进一步的研究。

对比分析在语言教学中有着广泛的应用,但也存在一定的局限性。下面扼要介绍其在教学中的作用及其局限性。

**3. 对比分析的作用与局限**

(1) 对比分析在教学中的主要作用

1) 对比分析可以帮助编制教学和考试大纲。

2) 在编写教材的过程中,对比分析可以帮助确定难点、重点,安排顺序,编制针对性练习,预防学生的错误。

3) 在教学过程中,对比分析可以帮助强调重点、补充练习,有的放矢地纠错。

4) 可以依靠对比分析选择测试项目,为不同语言背景的学生设计不同的考试。

5) 通过教学中的对比分析,可以发现只研究单语所发现不了的问题,为语言研究提供课题。

应该注意的是,对比分析只是语言教学的资源和背景,而不是教学的直接内容。因此,教学活动中不能大讲对比,特别是在国内的教学情况下学生语言背景复杂,对比分析的方法用得不好反而会产生负面影响。在国外语言背景单一的情况下,对比的方法更容易使用,也更有效。

(2) 对比分析存在的局限性

1) 对比分析的主要观点是两种语言的相同之处产生正迁移,不同之处产生负迁移,两种语言的差异越大,干扰越大,学习的困难就越大。实际情况并非如此。

2) 重视语言形式,忽视语言功能。语言形式上的对比分析不能解释学生功能、意义方面的错误,不能解释得体性方面的错误。

3) 重视语际干扰,无视语内干扰。事实上语内干扰所造成的错误比语际干扰所造成的错误也许更多更复杂。

4) 重在预防而忽视诊断。对比分析重在预测学习的难点和错误,而对

错误产生的原因和纠正措施不太关注。

## 二、偏误分析和中介语理论

### 1. 偏误分析产生的时代背景

对比分析的理论基础是结构主义语言学，而结构主义语言学的心理学基础是行为主义心理学。行为主义心理学认为人类的言语行为是可以分解、塑造的，塑造的主要方式是语言形式的机械操练，错误的言语习惯一旦形成就很难纠正，所以语言教学中要尽量避免错误的发生。这一理论曾一度支配美国并在很大程度上影响了别的国家的语言教学。20 世纪 50 年代末乔姆斯基的转换生成语言理论出现后，对行为主义心理学产生了巨大的冲击，同时也促进了认知心理学的发展。语言教学研究的重点由“教”转向“学”，把学习过程和特点作为研究的核心。认知心理学与行为主义心理学针锋相对的观点可概括为如下诸点：

1）行为主义心理学认为语言是一套习惯，是一个结构系统，而认知心理学则认为语言是一个转换生成的创造系统；

2）行为主义心理学认为语言可以通过反复的机械操练获得，而认知心理学则认为语言只能通过有意义的学习获得；

3）行为主义心理学主张尽量避免学习犯错误，而认知心理学则认为语言学习是一个不断假设、求证的过程，错误是不可避免的。

在认知心理学的冲击下，人们对以行为主义为基础的对比分析产生了怀疑，怀疑它的预测性和有效性。后来的研究发现学生的错误中受母语影响的不到 50%，那么，其余的错误是什么样的？又是怎样产生的？这些都是对比分析无法回答的。正是在这种情况下，偏误分析（error analysis，也译作“错误分析”）就因需要而产生了。

偏误分析就是对学生在第二语言学习中所产生的偏误进行分析，进而研究第二语言的学习过程。偏误分析是研究学习过程的捷径，也是研究学习过程的第一步。这也许就是偏误分析备受重视的原因。

### 2. 偏误分析的理论基础

偏误分析的理论基础是中介语（interlanguage）理论。中介语是第二语言习得中的一个重要概念，专指第二语言习得者所形成的与母语和目的语两者都不相同的个人独特的语言系统。Interlanguage 这个概念最早是由赛林格（L. Selinker）于 1969 年提出来的。对于这一不断离开母语向目的

语渐进的语言形式，有人称之为“个人特异方言”、“习得者语言”、“渐进系统”与“过渡语”。国内对 Interlanguage 的译名也不一致，除了“中介语”以外，还有“族际语”、“中继语”、“过渡语”等。关于中介语，至今尚缺乏全面的描述。根据现有的研究成果，中介语具有以下特点：

1）中介语是独立的语言系统，既不同于母语，也不同于目的语；

2）中介语是合法的语言系统，它有自己的规则，学习者不是随意地使用这种语言，而是有意识地创造性地使用这些规则；

3）中介语是不断变化的语言系统，逐步由母语向目的语靠近。

**3. 偏误的来源**

已有的偏误分析的成果认为偏误的来源甚多。大家所公认并有较多研究的有下列几种。

1）语际迁移(Interlingual transfer)。即母语知识向目的语迁移。

2）语内迁移(Intralingual transfer)。即第二语言学习过程中，目的语内部规则的相互迁移。语内的负迁移也叫过度泛化。泰勒(1975)发现，初级阶段语际迁移占优势，而随着学习者水平的提高，语际迁移越来越少，语内迁移则越来越多。偏误分析的一个主要贡献就是承认语际干扰之外的偏误来源——语内干扰。

3）学习语境(context of learning)。学习语境主要是指课堂、教材、教师、社会环境。教材或教师不恰当的解释和引导会导致理解的扭曲，学习者由此而产生偏误。

4）交际策略(communication strategies)。这种偏误与语际、语内迁移互有交叉。主要表现形式有：回避、预先制作的模式、认知和人格特点、求助于权威、语言转换等。

5）文化迁移(cultural transfer)。这类负迁移往往表现为语言形式上没什么问题，但是从目的语看不得体。

**4. 偏误的类型**

不少研究者从不同的角度对偏误进行了不同的分类。有人根据偏误的严重程度将其分为全局性偏误和局部性偏误。有人根据学习者能否意识到自己的偏误以及能否纠正并加以解释将偏误分为 3 个阶段：系统前阶段、系统阶段和系统后阶段。还有一种具有很强的逻辑推理性的分类，将偏误分为添加、遗漏、替代和错序 4 种。添加就是在语句中增添了不必要的成分，遗漏就是在语句中缺少了必要的成分，替代就是用同义词、近义词或短语代

替应该出现的词语(往往是以简单的代替复杂的),而错序则是语序不合乎目的语规则。

此外,偏误分析还应注意区分失误和偏误。失误是口误或笔误,是偶然现象,它不是系统的,不反映说话人的语言能力,一旦出现失误,说话人有能力改正。而偏误则是对正确语言的偏离,是成系统、有规律的,反映了说话人的语言能力。后者是偏误分析的对象,而前者则必须剔除出去。

5. **偏误分析对语言教学的作用**

偏误分析在第二语言教学中的作用是显而易见的。它可以使教师对学生的偏误有一个较全面较客观的认识,学生的偏误也为教学提供宝贵的反馈,偏误分析还是编写教材的重要依据。

最后,简单介绍一下国内的偏误分析和中介语的研究情况。我国最早引进中介语理论的是鲁健骥的《中介语理论与外国人学习汉语的语音偏误分析》(1984)。鲁文的发表引发了一系列对汉语偏误分析的研究,至今偏误分析和中介语理论的研究仍是对外汉语教学研究中的一个热点。但就研究方法而言,现有的研究成果中经验式研究较多,理论式研究较少;举例式研究较多,定量统计研究较少;归纳式研究较多,实验研究较少。当然这种状况不仅仅存在于偏误分析和中介语理论的研究之中,也许整个对外汉语教学界都应努力提高学术品位,这样才能真正和世界接轨,和世界对话。

## 参考书目

1. 〔加〕麦基. 语言教学分析. 北京语言学院出版社,1990
2. 高兰生等. 英语测试论. 广西教育出版社,1996
3. 李扬主编. 对外汉语教学课程研究. 北京语言文化大学出版社,1997
4. 刘镰力主编. 汉语水平测试研究. 北京语言文化大学出版社,1997
5. 刘珣主编. 对外汉语教学概论. 北京语言文化大学出版社,1997
6. 刘珣. 对外汉语教育学引论. 北京语言文化大学出版社,2000
7. 刘英林主编. 汉语水平考试研究(续集). 现代出版社,1994
8. 刘英林主编. 汉语水平考试研究. 现代出版社,1989
9. 盛炎. 语言教学原理. 重庆出版社,1990

10. 王建勤主编.汉语作为第二语言的习得研究.北京语言文化大学出版社,1997

11. 王敏敏.测试语言交际能力的汉语水平考试研究.上海师范大学硕士论文,1999

12. 肖奚强.现代汉语语法与对外汉语教学.学林出版社,2002

# 第三章 语 言 测 试

## 第一节 实验方法在语言学中的应用

### 一、定量方法和定性方法

#### 1. 定量方法的产生

现代语言学诞生的标志是结构主义语言学即描写语言学的出现，这种理论受美国心理学中的行为主义的影响很深。尽管 20 世纪 60 年代以后乔姆斯基的转换生成语法打破了结构语言学的一统天下的局面，出现了很多语言学流派，但仍然在不同程度上继承了结构语言学的方法：在研究中强调的是语言形式，关心的是语言的普遍规则，而对语言差异不感兴趣。功能语言学即社会语言学的出现拓展了语言学的视野，把语言学研究的对象延伸到了句子以外的领域。可以这么说，功能语言学对语言学的发展在某种程度上是方法论的发展。而近些年认知科学的发展，特别是认知心理学研究方法上的突破，使实验方法在语言学研究中越来越得到重视，成为现代语言学的分支学科如应用语言学、心理语言学、认知语言学等最常使用的研究方法。

定量方法也就是计量方法。科学的本质可以说就是数学。现代科学如力学、天文学、化学等在发展自己的理论时广泛地运用了数学这个工具，以一些公式来表示自己的定律；在社会科学和人文科学领域里，也引入了数学的方法，如经济计量学、历史计量学等等；在语言学研究中也已经使用了定量的方法，例如语料库语言学就是计量语言学的一个分支，其他诸如自然语言处理、机器翻译、语言教学、信息检索等方面，人们对定量方法的兴趣也越来越高，数学的方法已经成为语言研究中不可缺少的部分。

2. 定量方法和定性方法的差异

从方法论角度看，定量方法和定性方法可以说是两种完全不同的研究方法，有人甚至把这两种研究方法看成是不容易调和的“两种文化”，其实这两种方法所得出的结果往往互相补充。定量研究也可以称为实验性研究，定量方法也就是实验方法。定量方法和定性方法的区别可以归纳为下表：

| | | 定性方法 | 定量方法 |
|---|---|---|---|
| 研究基础 | | 现象学观点 | 逻辑实证主义观点 |
| | 1 | 强调亲身参与活动以获得经验 | 强调用实验方法来获取数据 |
| | 2 | 只有通过个人主观经验才能认识人类行为 | 只有摆脱主观状态才能了解社会现象的因果关系 |
| | 3 | 了解就是移情 | 了解要保持距离 |
| | 4 | 依赖定性数据 | 依赖定量数据 |
| 研究手段 | | 自然观察 | 操纵和控制 |
| | 1 | 观察面广，但分散 | 观察面窄，但集中 |
| | 2 | 变量不加控制，有利于了解它们的复杂关系，但容易顾此失彼 | 变量有所控制，有利于了解它们的因果关系，但容易简单化 |
| | 3 | 注意内容，但容易忽略形式 | 注意形式，但容易忽略内容 |
| | 4 | 解释力强，但容易主观 | 客观性强，但解释力弱 |
| | 5 | 接近现实，但时间长 | 时间短，但人为的成分大 |
| 研究方法 | | 归纳法 | 演绎法 |
| | 1 | 以数据为出发点 | 以假设为出发点 |
| | 2 | 没有事先形成的概念 | 进行预测 |
| | 3 | 可生成假设 | 检验假设 |
| | 4 | 研究成果：描述或假设 | 研究成果：理论 |
| 研究过程 | | 综合过程 | 分析过程 |
| | 1 | 从分体到整体 | 从整体到分体 |
| | 2 | 全面观 | 特殊观 |
| | 3 | 面向内部结构 | 面向外部结构 |
| | 4 | 了解过程 | 了解结果 |

(续表)

| | | 定 性 方 法 | 定 量 方 法 |
|---|---|---|---|
| 研究特点 | | 描写性特点 | 推断性特点 |
| | 1 | 无控制的自然观察 | 有控制的实验 |
| | 2 | 归纳与描写数据 | 归纳数据进行推断 |
| | 3 | 旨在发现形式 | 旨在验证假设 |
| | 4 | 效度高、信度低 | 信度高、效度低 |
| | 5 | 概括程度低:个案研究 | 概括程度高:多元观察 |
| | 6 | 强调动态性 | 强调稳定性 |

定量方法和定性方法都是科学的研究方法,在语言学研究中不能只局限于一种。一般来说,从事社会科学和人文科学的人比较熟悉定性方法,但在社会科学和人文科学中引进数学的方法更体现了当前科学发展中的文理渗透、学科渗透的特点,是有广阔的前景的。一些边缘语言科学的产生会使定量方法成为不可缺少的部分。而脱胎于定性研究的描述性研究,如结构语言学、功能语言学、人类语言学等不同的流派,它们采用的描述性研究方法实际上是一种介于定性研究和定量研究之间的方法。例如描述性研究虽然也强调自然观察,但它不像定性研究那样在调查前不带任何框框,不提出假设,而主张可以根据现存的数据或现象事先提出假设,作为考察的基础;描述性研究也主张归纳法,但不排除使用演绎的方法;描述性研究不像定性研究那样,在可能的范围内也使用定量的手段。

## 二、实验方法的意义和特点

### 1. 实验方法的意义

实验方法是建立在可计算性原则的基础上的,是把计算的理论、模型和人的行为加以比较的重要手段。实验方法又是建立在信息处理原则的基础上的,是把人脑看成是一个有限度容量的信息处理系统。

实验方法是科学研究经常使用的方法和程序。科学家认为,科学知识按照"前科学—常规科学—危机—科学革命—新的常规科学"的程序发展。什么叫常规科学?常规科学有两个主要的特征,一是它们是前所未有的,足以从相互竞争的科学活动方式中吸引一群持久的支持者;一是它们是开放

性的，足以提供各式各样的问题给不同组别的实践者去解决。具有这两个特征的科学成就称之为“范式”(paradigm)。从前科学到常规科学的标志就是范式的出现。上述这种科学思想证明了科学探索是一个不会终结的过程：不但在常规科学中，就是在形成常规科学以前或以后，都必须进行科学实验。科学实验主要涉及到以下两个问题：

1) 效度的问题，即怎样准确判断哪些事件和个案是相同的。为了进行效度的检验，我们必须精心地设计实验，保证参加实验的被试者有代表性，在他们身上所取得的结果不会是特殊的个案。用抽样的实验方法保证被试者的随机性，使相同的人都有同等机会成为被试者；用描写统计学的实验方法整理结果，对样本进行描写，以显示样本和总体的关系。

2) 信度的问题，即怎样把所得到的结果应用到更一般的场合。为了进行信度的检验，我们需要证明发现是有规律的，而不是偶然的。这就需要用推断统计学的实验方法，把实验重复做一遍。但要做到这一点很不容易，因为实验条件会改变，这需要我们考虑在多大的程度上能够获得相同的结果，结果的一致性，就是信度。

**2. 实验方法的特点**

实验方法有如下的特点：

1) 系统性(systematic)。在实验时必须有一套要遵守的程序性规则，这些规则包括怎样找出变量和定义变量，怎样设计实验以观察变量和决定变量的作用，如何把数据和原来的研究课题、假设联系起来等等。这些规则使实验系统化。

2) 逻辑性(logical)。实验方法所执行的规则和程序形成一种直截了当的、逻辑性强的型式，使研究一环接一环地展开，每一环和下一环都有逻辑关系。研究者可以根据内部效度的要求去检查实验中的每一个环节，也可以使用逻辑来检查所概括的外部效度。显然，研究效度的逻辑比靠观察数据，然后拍脑袋想出主意来要科学得多。

3) 经验性(empirical)。实验方法从现实世界里收集数据。数据的类型虽然很多，但它们都可以量化，都能够表示为一些可捉摸的数字。研究者可以靠处理这些数字从而和客观世界的研究联系起来。因此，实验方法是一种通过控制和操纵的经验性的研究。

4) 简约性(reductive)。研究者对数据进行分析时，往往会把纷繁的个别事件和个别对象简约为可理解的概念和范畴，应该舍去一些特殊性和独

特性,得到的是更为概括的事物间的相互关系,这一过程实际上就是对现实的内部规律进行抽象化。简约化使实验不仅是在描写事物,而且是在解释事物。

5)重复性和传递性(replicable and transmittable)。实验是可以重复的,这是因为实验的设计、数据的收集、统计和分析都有很大的透明度,别人可以重复这些过程来检验实验结论是否正确,也可以让别人使用研究的结果去从事另一项研究。

要指出的是,社会科学实验要比物理实验困难得多。这种困难表现为以下几个方面:

1)实验的对象是人,而人是很难重复的,他们不可能像从口袋里捡豆子那样容易作随机抽样;

2)社会科学的实验并非在实验室里做的,而是在现实的世界里做的,所以,实验室条件对所有的被试者不容易保持一致,也很难区别实验和实际活动;

3)人是一个十分复杂的统一体,即使是同一个被试者,在不同的外部环境下,每一次测试都会显示出智力上的、生理上的、心理上的差异,从而影响测试的结果。

**3. 实验方法在应用语言学、心理语言学中的应用**

但是,由于采取定性的研究方法更难作出实验和实际活动的区别,更不容易作出对人的复杂因素的分离工作,所以实验方法即定量方法在语言学的研究中越来越被广泛采用。从认知科学必须遵循的两条基本原则,即可计算性原则(所建立的理论必须在计算上有明确规定,在算法上可以执行,这些执行程序保证理论上提出的各种假设是明示的,为理论的检测提供方便;在输入改变后可以观察结果)和信息处理的原则(把人类看成一个有限度容量的信息处理系统,大脑在执行某项任务时所使用的机制和过程,可以理解为一项任务被分解为若干阶段,然后作有序处理的信息流程)出发,一般的实验方法有表征类型分析法、扣除法、递加因素法、信号监察法、计算机模拟法等等,这些方法被广泛运用于应用语言学的研究和心理语言学的研究上。

从应用语言学的研究来说,还有一个重要的问题是实验的设计。应用语言学(狭义的)是在语言学和教育学的边缘上产生的,它吸收了教育实验设计和教育测量的理论和方法,研究探讨外语教学的认知基础和过程、语言

习得、语言测试等问题。应用语言学的研究范围十分广阔，又有许多因素纠缠在一起，所以对实验的设计要求较高。从实验设计的程序来说，一般经过选择课题、提出假设、评论文献、决定变量、操纵和控制变量、观察和测量等过程。任何一个实验都要有严格的规程。

从心理语言学的研究来说，要使用许多心理测量的方法。心理语言学有基础研究，如语言和记忆、语言的理解和产生、语言和认知、语言和思维的关系等等；心理语言学也有应用研究，如语言的习得过程、语言学习者的个体差异、语言学习者的内部大纲等等。语言是可观察的，而心理状态是无法观察的，然而语言又难以控制，靠自然观察的方法往往费时失实，这就需要采取实验的方法。同样，心理测量和一般的实验方法相同，也要经过提出假设、决定变量、观察与测量、统计与分析等过程，也要有严格的程序。

## 三、测量的信度和效度等问题

### 1. 信度的测量

观察和测量是进行实验的一个环节，测量的内容指标主要是信度和效度。

(1) 信度公式

信度指的是测量的一致性。一个实验或测试的信度指它在重复测量时产生同样结果的程度。当测量的对象是人时，因为人的因素很不稳定，测量中的误差就在所难免。假设一次测量所得到的分数为 $X$，测试者真正的分数为 $t$，测试这所得的分数中包括的误差为 $e$，那么，这三者的关系可以用下列公式表示：

$$X = t + e \text{ 或 } t = X - e$$

但是信度的概念不是仅仅指对一个人某种能力的测量，而是指对不同人的反复测量，使它能表示不同的人的 $X$，$t$ 及 $e$。再假设用 VAR(方差)来表示它们的差异，那么：

$$\mathrm{VAR}(X) = \mathrm{VAR}(t) + \mathrm{VAR}(e)$$

如果我们在方程式的两边都除以 VAR($X$)：

$$1 = \mathrm{VAR}(t)/\mathrm{VAR}(X) + \mathrm{VAR}(e)/\mathrm{VAR}(X)$$

这样，信度($r$)可以表示为 VAR($t$)和 VAR($e$)的比率：

$$r = \mathrm{VAR}(t)/\mathrm{VAR}(X) \text{ 或 } 1 - \mathrm{VAR}(e)/\mathrm{VAR}(X)$$

当信度 $r$ 的值为 0.50 或 0.75 时，误差过大，一般的测量均要求信度在 0.90 以上。

(2) 信度估算的方法

信度估算的方法很多，一般有如下几种：

1) 再测法：用同一套试题在不同的时候测试同一批学生，然后在求出两次测试的相关系数，如果两次结果完全相同，则信度为 1.00。这种测试方法的问题在于：① 测试者的水平会不可避免地发生变化；② 一次测试结束之后，测试者会记住一部分题目和答案；③ 实际操作时，往往不容易找到原有的测试者。

2) 平行试题法：用两套平行试题代替同一套试题，以避免测试者的机械重复。平行试题在格式和内容上应保持一致，其他方面与再测法相同，只是要求两次测试的时间不能相隔太远。这种方法的主要问题是很难出两套真正平行的试题。

上述两种方法无法得知一个人的真正的分数，因此也无法得知信度，只能将相关系数作为估计的基础。如果从误差源的角度来考虑计算测试分数的信度则可采取下面的方法：

3) 对半法：把试题分为两半，然后决定这两半的分数在多大程度上是一致的，这是以下列条件为前提的，即假设这两半是等值的，它们的平均分和标准差是相等的，而且是相互独立的。这种方法的问题是在大多数情况下，不能用随机的方法来进行对分，且试题中的不少题目难以互相独立。

(3) 阅卷员的信度问题

阅卷员在评阅主观性题目时（如作文、口试等），常常会有误差，这就牵涉到阅卷员的信度问题。

1) 阅卷员的内部信度(intra-rater reliability)问题：一个阅卷员评阅主观性题目时，必须使用一定的评分标准。要了解这个阅卷员的评分信度，可以让他在不同的时候评阅同一份试卷，即随机地抽出一些试卷进行重评，然后计算两次评分的相关系数。

2) 阅卷员之间的信度(inter-rater reliability)的问题：不同的阅卷员掌握的评分标准在评阅主观性试题时，不容易取得一致。要了解他们之间的信度，一般是让两个阅卷员评阅同一份试卷，然后计算其相关系数；或者把试卷随机分配给阅卷员，然后把每个人的评阅分数的分布和整体评阅分数

的分布加以比较,并根据差异进行调整。

(4) 影响信度的因素

影响信度的因素大致有如下几项:

1) 试题的长度:试题长,项目多,信度就会提高。

2) 试题的同质性:如果一份试题考的都是同一能力倾向,信度也会提高。

3) 试题的区分度:区分能力强的项目越多,信度也会越高。

4) 考生的差异性:如果考生能力差异很大(表现为标准差很大),信度也就很大。

5) 考试时间的充裕性:目前所使用的信度公式,是以考生能够做完所有的题目为前提的,考生时间充裕,用于估算能力考试的信度就高。

**2. 效度的测量**

(1) 效度与信度的关系

效度指的是我们所获得的证据在多大程度上支持我们根据分数所作出的推断。效度和信度是互为补充的。考察信度是为了回答这样的问题:"测试结果中有多少方差是测量误差引起的?""有多少方差是测量误差以外的因素引起的?"而考察效度是为了回答这样的问题:"有哪些能力可以说明测试结果中的可信方差?"所谓的"可信方差"是指测量误差以外的因素所引起的误差。因此我们可以说,信度考虑的是测试结果中有多少方差是可信方差,也就是说,信度是用相似方法去测量同一倾向的一致性;效度考虑的是哪些能力会导致可信方差,也就是说,效度是使用不同的方法测量同一倾向的一致性。

(2) 内容效度(content validity)的问题

测试的目的是为了了解测试者在一些实际的环境下是如何运作的。但是我们不可能在测试中再现所有的环境,而只能选择一些样本,如何根据测试者在这些样本中的表现,评估其运作的能力。如果测试的样本能够充分地代表总体,那么测试在内容上就是有效的。内容效度涉及两个方面:一是内容的关联性,这牵涉到规定的能力范围、测试这些能力的方法及测试者作出的相应的反应;另一是内容的覆盖范围,即测试者所完成的任务在多大程度上反映能力范围的要求,这牵涉到测试项目的代表性、测试项目的数量等问题。

(3) 预测效度(predictive validity)和共时效度(concurrent validity)的

问题

这两种效度都属于与标准有关的效度，即表示测试结果和某些标志测试者能力的标准的关系。预测效度表示的是某些标志测试者将来的能力的标准，例如根据摸底考进行分班，再和测试者以后的考试分数进行比较，就可以了解摸底考的预测效度。后者表示的是一个测试结果与另一个同时使用的标准的结果的关系，例如要检测 HSK 模拟题的标准化的程度，可以让测试者在差不多相同的时候参加正式的 HSK，以此了解模拟题的共时效度。

(4) 构想效度(construct validity)的问题

构想效度要考察一个测试者在多大程度上和我们根据某一理论所作出的预测相一致，它验证的是我们所作的假设是否有效。构想效度是一个把其他的效度统一起来的概念，在验证构想效度时，通常使用相关的方法，即考虑一些相关的测试题目的互相影响和互相渗透，或者使用多维度的分析方法，即考察各个相关的聚合(convergence)和区别(discrimination)的型式：聚合是指共时的标准，考察的是使用不同方法测量同一倾向的一致性程度；区别是指使用不同的或相同的方法，测量不同的倾向产生不同结果的程度。

## 第二节　现代语言测试的理论框架

### 一、语言测试在语言教学中的地位

#### 1. 语言测试对应用语言学的贡献

语言测试在语言教学中处于中心地位，是从属于应用语言学这一学科的。过去语言测试在应用语言学中只具有边缘的地位，在 30 多年的发展中，现在已成为具有主导地位的学科了，这是因为它通过运用各种可行的方法，处理语言学习和实际运用的许多问题，已从单纯的实用阶段发展到理论与实践相结合的阶段。这样，语言测试已经能够在学术上占有一席之地，并逐渐发展了自身的解释能力。语言测试为语言教学提供了目标，并对教师和学生成功地实现这些目标进行监控，而且为语言的讲授和学习提供了试验和调查的方法。语言测试对应用语言学的贡献可以归结为 3 点：

1) 使应用语言学的理论框架转为实际运用；

2）成为为教学进程和教学大纲确立的目标和标准；

3）为应用语言学的研究提供了方法论上的借鉴。

语言测试的影响之大，作用之突出，与语言理论和语言实践的结合之紧密，确实是值得我们重视和研究的。然而，以往的有关语言的各方面著作中，对语言测试的阐述都极为简单，造成这种现象的原因，在一定程度上是由于怀疑对语言的各种定量分析的方法，同时也因为过分相信充满人类智慧的文化知识不能简化为纯粹的数字的结果。

**2. 语言测试所涉及的语言理论和语言实践的问题**

语言测试需要考虑理论与实践方面诸多因素：从理论方面来说，要考虑语言观和语言学习观；从实践方面来说，要考虑内容、实施和评估。不同的语言观决定语言测试不同的方法、内容和题型。从语言测试的发展看，经历了3个阶段的发展：第一个阶段是和传统的第二语言教学紧紧联系在一起的，一直延续到20世纪40年代，可以称为心理测量-结构主义语言测试阶段，其语言学的理论根据是结构主义语言学；第二个阶段形成并占主导地位始于20世纪40年代，这一阶段的理论体系直到八九十年代初，仍是世界范围内测试领域里最大的主流，可以称为心理语言学-社会语言学语言测试阶段，十分重视“总体综合法”的研究；第三个阶段崛起于60年代后期，80年代初，第三代测试理论虽然已占上风，但在实践上仍和第二代模式平分秋色，至80年代后半叶，开始成为与第二代主流抗衡的强大潮流，这一阶段可以称为交际法语言测试阶段，受到社会语言学家提出的“语言交际能力”概念的深刻影响。下表反映了语言测试所涉及的理论、实践等问题在这3个阶段的不同表现：

| | | | 第一阶段 | 第二阶段 | 第三阶段 |
|---|---|---|---|---|---|
| 理论 | 语言观 | 内涵 | 教的、考的应该是知识 | 教的、考的应该是技能 | 教的、考的应该是能力 |
| | | 外延 | 由于教的、考的是语音、词汇、语法知识，所以语言观的外延主要限于语音的形式系统 | 教的、考的严格限于语言形式系统，外延基本上与第一阶段相同 | 交际能力包括语言能力，即语言形式结构系统本身的操作能力；语篇能力，即语言结构形式置于语篇中运用的能力；语用能力，即语言结构形式置于情景中运用的能力 |

（续表）

| | | | 第一阶段 | 第二阶段 | 第三阶段 |
|---|---|---|---|---|---|
| 理论 | 语言学习观 | | 把语言学习看作知识的接受；得到的知识主要是陈述性知识，知识的接受主要靠背熟记牢，语言测试就是考记牢的知识 | 把语言学习看作刺激反应的技能训练：技能训练不需要认知，只是操作语言形式的技能，可以一项一项地学，语言测试以考这种脱离意义和运用的技能为主 | 是潜意识和认知性的学习：交际能力主要靠习得而获得，习得通过交际运用，所以语言测试应通过交际运用；语言学习必须以学习者为主体，测试应给受试者更多的主动性 |
| 实践 | 内容 | 具体化 | 语言点：语音、词汇、语法及惯用语的知识 | 有限数量的搭配力最强、使用率最高、概括性最广的基本形式结构，重点放在句型上 | 交际能力的体现是交际行为，交际行为是测试内容的具体化 |
| 实践 | 内容 | 选择 | 以命题人和教师的经验和主观判断为依据 | 除对目标语作结构分析外，还作目标语与母语的对比研究，使用频率的统计作为依据 | 要进行学习者需要的分析，需要就是教学和测试的目标行为，是为了给无限的交际行为限定一个范围 |
| 实践 | 内容 | 说明 | 教学内容缺乏质和量的明白说明，测试内容的说明根本不存在 | 对目标语结构的形式、规律和系统作客观的分析研究并加以说明，测试具有明晰性 | 从定量、定性两方面入手，采用分维说明是为了使多维的、不定型的行为具体化，所以说明难以达到高度的明晰性 |
| 实践 | 实施 | 方式 | 知识型试题：大部分要求"记"：形式、规则、目标语与母语的对应；另有一部分要求用这些记住的知识，如造句、改写、作文等 | 着重测试对语言形式的反应及操作技能，常用语法填空、词汇填空、句型变换等试题，还推出多项选择题（MC 题型）和是非题（T/F 题型） | 教学上的练习，口语称为事型练习，试题是事型试题，设定一个情景，然后在这一情景下，完成一系列听说读写的测试 |
| 实践 | 实施 | 考点 | 主要是语音、词汇、语法的知识点，重读写，轻听说 | 理论上认为每题只考一个语言点，强调口语，注意给听说技能（特别是听）一定的比重 | 对听说读写作综合处理 |

（续表）

| | | | 第一阶段 | 第二阶段 | 第三阶段 |
|---|---|---|---|---|---|
| 实践 | 实施 | 语境 | 较“死”的试题中没有语境，较“活”的试题中可能有，但往往不真实 | 强调考点不超出语言自身，一般只考到句子，语言缺乏语境 | 强调语言必须放在语境和真实的交际情景中去考 |
| | | 真实性 | 多数选用完全属于书面语的语言材料 | 较注重口语，比第一阶段具有真实性 | 具有真实性，语言必须适合特定的时、地、人物身份、题材体裁、交际手段和目的等 |
| | | 焦点 | 以语码为焦点，要求测试者把注意力放在语言形式上 | 以语码为焦点 | 以信息为焦点 |
| | 评估 | 评估标准 | 主要标准是准确性，没有意识到流利是同样的语言属性，更没有考虑到得体性的问题 | 更强调语言形式的准确性，有绝对化的倾向；重视流利，但认为流利单纯是语言形式结构操作的熟练问题；完全忽略得体性这个评估标准 | 评估对象是交际行为，评估标准是交际的有效性，交际有效性把准确性、流利和得体性都包括进去，根据具体交际行为决定对准确性、流利和得体性的要求 |
| | | 判断方法 | 主观的量化评估，而且是两分法评估，对语言表现只作正确与不正确两种判断 | 由于采用了大量的离散项目测试题型，因此发展成为高度客观的两分法量化评估 | 有效性本身是一个质的概念、连续的概念：评估强调有效交际，不可能是量的评估；测试的标准不是正确与错误的问题，而是要看测试者用目标语做得怎么样的问题 |

## 二、语言测试的性质

语言测试的定义是什么呢？完整精确的定义应该是：语言测试，是对一组言语行为进行测量，这组行为是表征目标能力的全部行为的一个抽样。测试的结果，是为了可以对目标能力作出评估。

1．*语言测试是一门科学，是多门科学结合的产物*

语言测试的发展，接受了多门学科的馈入，但是归纳起来，馈入主要来自两个领域：一个领域给语言测试馈入内容，解决测试什么的问题；一个领

域给语言测试馈入手段,解决怎样测试的问题。

给语言测试馈入内容的学科,都可以归入内容科学(content sciences)里。这些学科又可以分成以下几类: ① 研究语言本身系统的学科,如语言学以及它的三大核心学科,即音系学、语法学和语义学等; ② 研究语言的性质、运用及学习的学科,如语言哲学、语用学、篇章语言学、语料库语言学、语体学等; ③ 语言学科和其他学科相结合后产生的边缘学科,如社会语言学、心理语言学、神经语言学、人类语言学、语言交际学等; ④ 虽然不直接和语言学结合,但对语言学有深刻影响,并有可能起到决定性作用的学科,如哲学、心理学、社会学、认知科学、信息科学等。给语言测试馈入内容,并不仅仅意味着语言测试能够直接从这些学科中取得内容,更重要的是这些学科可以提供语言测试对内容的性质、范围、层次、维度、抽样、实现等问题作出决定的原则、依据、理由和目的。

给语言测试馈入手段的学科,都可归入方法科学(methodology sciences)里。这些学科包括测量科学、实验方法论和信息技术。最基本的是统计学、教育测量学、心理计量学以及辅助语言教学和语言测试的视听技术、辅助语言教学和语言测试的计算机技术和方法、计算语言学、人工智能的应用等等。提供手段,不仅仅是测试的实施手段,还应该包括测试的设计和生产手段,测试结果的取得、计算、分析和解释手段,对影响测试本身质量的各种因素(有效性、可靠性、区分度、稳定性、适宜性、偏颇性)的检验手段、保证手段和消除手段,以及有关测试反馈的采集手段、研究手段和分析手段等等。

语言测试科学实际上就是介于上述内容科学和方法科学之间,它从内容科学里取得内容,从方法科学里取得手段,互相结合而成,所以也可以这么说,语言测试科学既是内容科学,又是方法科学。然而,语言测试科学与内容科学、方法科学之间却并非是单向的关系,而是双向的,互有联系的:一方面,内容科学、方法科学分别给语言测试科学提供内容和手段,另一方面,语言测试也会成为内容科学的实验手段,从这个角度说,语言测试其实也是一门方法学,只不过它是用于语言这个对象的方法学罢了。因此,语言测试科学的发展,要靠内容科学和方法科学发展的推进;语言测试科学发展的速度,是由给它馈入内容和手段的科学的发展速度决定的。

**2. 语言测试所包含的信息**

语言测试是从长期的教学实践和学习需要发展而来的。近些年来,已

用作探测和调查语言能力的系统方法。语言测试的实践历史，在其发展过程中往往以考试的形式体现，通过试卷，对语言应用和语言知识进行定量的专门研究。语言测试应包括两方面的内容，一个是语言方面，一个是测试方面。因此，语言测试需要考虑理论和实践方面的诸多因素，将密切关注着以下3个方面的信息。

1）关于语言技能方面的信息。语言测试是基于教育测试基础上的，而教育测试本身则是由测量学理论、心理学以及心理测量学的实践发展形成的。教育学中所运用的启发式方法，也同样适合于运用测试的某些领域。运用测试要了解学习者在何种程度上可以达到所要求的水平，以及在达到更具体的和更特殊的所要求水平的过程中，语言测试所占的地位。

2）关于语言发展方面的信息。这种信息对于语言学习者的发展具有明显的作用。当然，这里的学习者是指第二语言的学习者，而对于母语的学习者来说，作用就不会那么明显。语言水平应该是有层次的，学习者所选的学习内容也应该有层次。这方面的信息，通过评估，还能反映出学习者的心理语言学和社会语言学方面的情况。

3）关于语言知识方面的信息。语言使用者，无论是操母语的，还是操非母语的，他们之间的差别主要体现在他们对语言的熟悉程度上，具体表现在对语言表达的可接受范围所作的判断，以及他们语感的敏锐程度上，也就是他们对语言知识运用的程度。语言知识的信息对语言技能的信息会有影响，因为在技能中包含了知识的因素。

**3. 语言测试的目的**

语言测试应该达到什么目的呢？这可以归结为以下几点：

1）语言测试可以促进语言研究和语言教学。从语言研究来说，语言测试可以用来测试对语言理解和语言学习相关的假设，语言水平的程度，语言能力的结构及语言习得的自然顺序等问题，从而成为语言研究的一种方法和手段；从语言教学来说，科学的、符合教学原理的测试，有诊断和指导教学的作用，可以增强学生的学习动力，使教师和学生在学习过程中有阶段感、程序感和成功感。测试结果所说明的问题及其意义，常被用来作为对教学大纲和教学情况的检验。

2）语言测试可以用来选拔人才。人才的要素是文化素质，文化素质主要指思想深度和获取知识的技能。思想深度要靠语言来表达，知识的获取

也要靠语言这一媒介来进行。语言水平越高,获得知识越快,表达就越精确;获取知识的渠道越多,思想也越丰富。因此,语言测试常常被用来作为选拔人才的方法。也可以这么说,以效度为主的测试的运用过程就其本身而言,在很大程度上是为选拔目的服务的。

3）语言测试关系到对学习者、课程、方法等方面的评估问题。测试是评估的一种手段。首先可以通过测试了解学习者是否学到了应学的内容,例如第二语言的学习中有多少信息是课堂里得到的,有多少信息的得到是学习环境之外的,测试可以评估教学内容对交际能力习得的影响程度;其次可以了解课程的合理性、有效性等问题,例如第二语言的学习从情景大纲转到交际大纲,通过语言测试得出一系列数据,从而评估出何种大纲更为合适。另外,测试还常被用来作评估教学方法、教学效果的手段。

## 三、语言测试的两大根本要求——信度和效度

### 1. 语言测试中信度与效度之间的矛盾

这里主要讨论在语言测试中信度与效度之间的矛盾及矛盾的解决方法。我们认为,语言测试的信度和效度具有相互排斥作用,这是由信度和效度的本质决定的,更是由语言的本质决定的。信度与效度的矛盾体现在以下几个方面:

1）量和质的矛盾。量的测量容易保持高信度,但是语言测试要实现效度,却要求质的测量。量的测量,测量的对象必须可数,而我们测量的语言,不光是语言形式结构,而是不可数不可见的语言运用能力。能力可以表征为行为,虽然可见了,但仍不可数,从本质上来说不能用数字加数量单位表示。因此,如果偏重信度,可以把测试尽量设计为定量型的,但忽略了效度;反之,把测试设计为定性型的,又忽略了信度。

2）两分判断和连续判断的矛盾。信度要求两分判断,要求测试中的每一个测试点,都只有对或者错两种判断;语言测试的效度,却要求连续判断,因为语言运用,从得体性、流利、有效性,甚至准确性来说,都是一个程度问题,而不是单纯的是与非的问题。但是,如果完全按语言运用的原本性质去对它作连续判断,虽有利于效度,但有损于信度。

3）分割法与整体法的矛盾。分割法测试,符合信度的要求,因为这种方法把复杂的整体分割为各个性质单一的元件,元件小,数量多,有利于取

得客观、明确的数据，测试项目多，信度也就高了。但分割法违反了语言整体观：运用语言，不可能把单词、句子一个个分割开来，不可能把语法点、词汇点等一项一项分割开。分割开的一个个元件，已经失去了语言的基本性质。因此，从效度考虑，应该用整体法，但整体法的语言测试，不易取得客观、明确的数据，也无法有足够的测试项目，不容易取得高信度。

4）单维观和多维观的矛盾。真实的语言是多维的：语言的运用要牵涉到语码（包括语音、词汇、语法等）、传播渠道（包括听、说、读、写、听说、读说、写说、听写、读写等）、意义（包括定指和泛指、局部义和总体义、字面义和隐喻义、引申义和比喻义等）、语体（包括口语体和书面语体、叙述体和抒情体等）、环境（包括语篇语境、人、地、时、目的、已知信息等）、背景（包括社会文化乃至历史民族、传统等）等等维度，倘若从多种维度的角度去测试运用中的语言，达到了高效度，但无法达到高信度。

5）静态测试和动态测试的矛盾。如果把语言作为一个凝固的体系去考察，把人对语言的掌握，看作已成的结果去测量，这种测试称作成果测试或静态测试，它关心的是结果；如果把语言作为一个活的、动的、不断变化的体系去考察，把人对语言的学习与运用看作一个过程去加以测量，这种测试称作为过程测试或动态测试，它关心的是过程。成果测试信度高，效度低；过程测试效度高，信度低。

6）有限的语法和无限的运用的矛盾。一种语言的语法总是有限的，以它作为测试的目标能力，容易保证信度，但语言测试不能仅以语法作为测试的目标能力；语言运用能力是无限的，以它作为测试的目标能力，就很难保证信度，但无限的运用如何去取得客观的数据呢？如何去把握呢？

7）抽象的行为和具体的行为的矛盾。行为是目标能力的表征。如果把言语行为置身于没有真实语境的情况下，这种行为是抽象的。语法规则越抽象，它的概括性就越强；概括性强，外推性也强；抽样的外推性强，语言测试的信度就高。如果把言语行为置身于具有真实语境的情况下，这种行为是具体的。行为越具体，其作为抽样的外推性就越欠缺；抽样的外推性弱，信度也难以落实，但效度显然就要高。

*2. 信度与效度之间矛盾的解决方法*

信度和效度存在着以上 7 个方面的矛盾，对这种矛盾的处理，有四种不同的处理方法：

| | 特　点 | 存在的问题 |
|---|---|---|
| 1 | 保证高信度,忽略效度 | 在对语言测试的效度的理解进入新的高度的情况下,这种处理方法有逆于科学的进步方向 |
| 2 | 追求高效度,放弃信度 | 交际测试学派的观点,是测试学界的理想主义者,在目前情况下,是无法广泛实施,也很难得到普遍接受 |
| 3 | 消极妥协 | 认为信度、效度不可兼得,两者不必强求,能做到多少算多少,实际上不愿意作科学上的探索 |
| 4 | 积极平衡:用信度高的形式测试效度高的内容 | 不满足现有的测试程序,解决办法仍在探索之中,依赖内容科学和方法科学的发展 |

## 第三节　语言测试中试题的生产程序

### 一、试题的设计和制作

**1. 语言测试工作是一个系统工程**

大规模的语言测试工作是一个系统工程:横向系统,体现在它的试题结构上,这是测试的蓝图;纵向系统,体现在它的工作程序上,这是测试的流程图。流程图的制定,是为了保证蓝图的实现:首先要明确测试工作的步骤,有多少个环节,定出一个运转时间表;其次要明确人力分工,三方面专业人员(考务行政人员、统计分析人员、学科业务人员)如何分工合作,定出一个组织计划;最后要明确要求,每一项质量、数量指标,每一个具体规定都要清楚,定出一系列的规则和制度。纵向的系统主要用在大规模的语言测试工作中,适宜的设计和制作主要涉及到横向的系统,即有关语言测试中试题的生产诸问题。

**2. 试题的整体设计制作**

试题的设计制作从整体上说,应从测试本身内容和测试技术形式两方面考虑。从测试本身内容方面说,要顾及到三方面的问题:① 覆盖面。根据教学大纲分析确定测试范围,根据这个范围划分出测试的几个主要方面,试题的结构可以按照这几个主要方面再划定几个部分。例如 HSK 中级的试题结构分成听力、语法、阅读和综合知识四部分。② 比重。决定了覆盖面之后,还要决定每个部分的比重,各部分比重的确定,主要从测试的性质

出发，而不能只是简单地把分数凑成整数。HSK（初中等）中听力占29.4%，语法占17.7%，阅读理解占29.4%，综合知识占23.5%。③ 题量。题量主要和时间有关系。语言测试可以有速度测试(speed tests)和力度测试(power tests)两种方法，前者提供超量的试题，主要测量在规定时间里每人能完成的量；后者给的题量估计在规定的时间内是可以完成的，旨在测试能力而不是速度，现在国内的测试大都是后者。从信度的角度考虑，每套试题在合理时间的限制下，题量应尽量多。HSK(初中等)中听力理解50题，语法结构30题，阅读理解50题，综合填空40题。

从测试技术形式方面说，应着重考虑两个问题：① 技术安排。主要考虑试卷中各部分几个大题的先后安排，这要牵涉到考务实施的问题，特别是大规模的测试时分卷、评分、统计等，需要流水作业；还要牵涉到测试者的应试心理问题，一般来说，往往把内容较容易的放在前面，先易后难，把客观题如填充、选择、判断等放在前面，然后再是主观倾向较浓的问答题。这种安排，有利于测试者水平的发挥。② 版面设计。一般试卷分试题卷和答题卷，有些测试还提供材料卷等。试卷的版面如何设计，如何分页，如何装订，如何加封条等等，都必须从有利于考务，有利于答题，有利于把结果输入计算机，有利于机器阅卷或人工阅卷等等方面去考虑。例如分页考虑不周，就会给测试者带来前页后页翻来翻去的麻烦，甚至会有漏答题、错位答题的情况发生。

3. **试题的定型**

试题的整体设计制作的问题考虑以后，对于要多少大题，每道大题考什么已作决定，下面的问题就是考虑如何为每道大题定型，即考虑怎么考的问题。怎么考的问题，不单纯是题型的事，可以从3个方面作选择。

| | 客 观 题 | 主 观 题 |
|---|---|---|
| 特点 | 1. 每题只允许一个形式固定的答案<br>2. 适合于测试知识、测试技能<br>3. 如果测试运用，只能是识别性或接受性的运用 | 1. 答案不能限定只有一个形式<br>2. 适合于测试语言的口头产生性技能，以及口头、笔头的产生性运用 |
| 优点 | 1. 可以排除阅卷员之间，阅卷员自身前后评分的不一致<br>2. 从经济和时间上都提高效率，适合于机器评分 | 1. 能够测试语言运用的能力<br>2. 效度高，拓宽了测试的内容 |

（续表）

| | 离散考点试题 | 综合考点试题 |
|---|---|---|
| 特点 | 1. 每题针对一个语言点<br>2. 所用的题型，多为单词或单句的填空、改错、变换等 | 1. 把多个考点综合在一起，同时测试多个语言点或多个语言层次<br>2. 所用的题型有听写、理解、写作、翻译等 |
| 优点 | 1. 测试什么十分明确<br>2. 评分能做到客观，题量可以扩大 | 1. 可以同时测试多个语言点<br>2. 能够运用听说读写综合技能 |
| | 无语境试题 | 有语境试题 |
| 特点 | 试题中的目标语以分散的单词、分散的单句等形式出现，不超出单独的句子这个层次 | 试题中的目标语是一段连贯的话或者是完整的语篇 |
| 优点 | 信度高，可行性强 | 效度高 |

总之，试题的设计制作，无论从整体上说，还是从每一道具体的题目来说，都必须以可行性为基础，以效度和信度为目标，整个测试的 3 个度，最终在具体的试题上得以落实。

## 二、试题的评分

### 1. 客观题的评分

评分工作是关系到测试信度的最重要的一个环节。评分工作分为主观题评分和客观题评分两大部分。客观题的评分相对主观题的评分要容易一些。就目前国内的情况来说，有三种评分方式：① 人工评分；② 人工输入＋机器评分；③ 机器读入＋机器评分。第②种方式、第③种方式就效率和效果来说，都要远远优越于第①种方式。从一项统计资料看，从效率上说，第②种方式比第①种方式省时间，省人力，试题的复用率也高，而第③种方式在上述 3 项指标中，非但比第①种方式效率高，比第②种方式也要高得多。特别从试题的复用率看，第①种方式试题测试后无法保密，失去入库再用的可能性，对试题生产劳动是极大的浪费。从效果上说，用第①种方式评分，而且在集中评分的情况下，误评试卷率可达 20%左右，而用第②种方式，误评试卷率就马上下降至 1%～2%之间，用第③种方式评分，误评试卷率可下降到零。因此，大规模的测试，应尽可能地采用第②种、第③种方式。

### 2. 主观题的评分

主观题的评分，主要针对产生性技能和产生性运用的试题而言的。这

一类的试题,如口试中的朗读,笔试中的作文、翻译等等。主观题的评分,关键要设法排除 3 个不一致,即评分场所的不一致,评分员之间的不一致及评分员自身的不一致。为了排除这 3 个不一致,就要采取一定的措施,限制一定的条件,这些措施和条件大约有以下几个方面:

1) 制定统一的评分标准。有两种方式:综合性评分(global marking)和分解性评分(analytical marking)。综合性评分的指导思想是把语言放在交际运用中作为一个整体去评价的,文章应该通篇考虑,评分是从整体去判断其要表示的意思。评分标准的具体内容,正是要描写区分达到不同交际目的的不同程度,并规定每一种程度的应得的分数。分解性评分的出发点是把整体分解为部分,各个部分独立评定给分,最后再加成总分。分解有很多方式,如评判作文时,可以把形式和意义分解开,形式下可以分成句子通顺、用词准确等等,意义下可以分成内容完整、主题突出等等,每一部分给相应的分数。上述两种评分方式,都有利弊:前者符合语言测试体系发展的方向,但难以控制评分的标准;后者不符合现代语言教学和测试的新观点,但较易控制评分的标准。目前国内大规模的语言测试,都采用半综合半分解的路子,既可取得信度,也能保证一定的效度。

2) 培训评分员。有了统一的评分标准,没有一批经过培训的评分员,是无法起到应有的控制作用的。尤其是综合性的评分标准,由于标准本身只能定得比较抽象,评分员必须通过培训才能取得一致的认识。即便是分解性的评分标准,评分员的培训也是十分必要的,因为再细的标准也不可能顾及到所有的可能性。培训一般要经过理论认识、对本次测试试题的认识和实际试评等步骤。作为评分员,应该服从统一的评分原则,避免造成对受试者的不公平。在评分的过程中,如果两个评分员同时评一份卷子或者一个受试者,两人给分的差距太大的话,可以请同样经过培训的第三位评分员为之评分,三个分中取两个相近者,这也是保证评分公正的一个措施。

3) 改善评分方式。无控制的单人评分,仍然是目前用得最多的主观题评分方式,但是这种评分方式得出的结果往往是不可靠的,加以控制的单人评分,评分的结果虽然比无控制的要可靠一些,但要通过计算机来调整分数,这是比较麻烦的事。那么,双人评分是否可以说解决了问题了呢?也不完全如此。要使双人评分的结果尽可能可靠,必须做到以下几点:第一,评分员必须经过统一的评分培训,培训的内容必须一致;第二,评分地点虽然是集中的,但要求两位评分员独立评分,互相不知道对方所给的分数,然后

取分数的平均值;第三,受试者的情况和身份不被其中的任何一位评分员所知,以避免评分员在认识受试者的情况下产生评分上的偏颇;第四,当两人在评分上产生较大差异时,必须请第三位评分员重新评分。当然,最理想的方式是机器评分。但这种评分方式需要解决不少问题:如受试者的试卷输入机器的问题,高级光电字符阅读器的制作问题等,更重要的是要解决人工智能对试卷作出判断并按要求评分的问题。如果这些问题都解决了,那么机器评分就会消除主观性评分的3个不一致性,从根本上解决了主观题信度的问题。

## 三、试题的质量控制和统计分析

### 1. 试题的质量控制

试题生产过程中的质量控制,可以通过多种方式进行,通常采用以下几种手段:

1) 建立审题员制度。试题的数据制作往往是个人独立完成的,但个人制作出来的试题,不能再由个人去检验。因为个人制作出来的试题,须经过更多的眼睛从尽量多的角度去衡量判断,才能尽可能地发现问题,发现不可接受之处。审题还须由多人集体去做,在审题前,要让每个人有充分时间审读和考虑要审的材料和试题。审题过程必须尽量充分讨论,尽量做到全体同意,审题组里即使有一人对某一道题有异议,也要考虑到他的想法意味着有人会从这个角度去看这道试题,到正式测试时,可能会有数十人、数百人与他有同样的想法,这就会影响到试题的效度和信度。

2) 实施试测、试做方法。在大规模的语言测试以前,最好通过实践来检验试题的质量。实践有两种方式,一种是试测,一种是试做。试测是向题库输送试题之前必不可少的质量控制程序,而试测后的题目分析,又是试题生产中审题改题的重要依据。然而,试测是正式测试的"彩排",要求各种条件和实际测试时尽量相似,测试的对象在类型上和水平上还应该和实际测试时的受试群尽量一致,所以,在实施上会有一定的困难。试做则是比较简单的试验手段,不必找到足够数量的和实际测试时受试者相同水平的对象,只要找两三位(至少是两位)受试者,在不知道答案的情况下,把整套审题做一遍,以此来发现可能存在的任何问题:如时间不够,客观题答案不对,甚至打印、装订、编号上的错误。试做费时不多,但也能发现一些问题,有利于事先排除一些可能会影响效果的干扰。

3）推行表格分析方法。在试题的生产过程中，有一些项目会用表格注明，在进行质量控制时，可以利用这些表格进行分析。例如，在客观性试题设计制作前，命题员会对一些题（如语法题和语音题）应测试的语言点项目用表格的方式列出。这类试题的考点，一般也比较明确，往往是一道试题考一个或两三个语言点。对这类试题的质量控制，只要了解每道题是考哪个语言点，分析时按原有的表格归类，大致可了解试题与设计方案之间的异同。主观题的设计制作前，对测试应覆盖的功能范围应已明确，以功能为考点的试题，一般只限于产生性技能的试题，主要是听和写的试题。功能分析应以测试的水平而定，水平较低的测试，可以分析到微观功能，也就是说可以将功能尽量量化，量化成一个个详细的考点，做成表格；而水平很高的测试，则只要分析到宏观功能就可以了。以测试微观功能为主的质量控制，可以对照表格进行分析；宏观功能的测试，无法分解为详细的考点，只能在控制试题总体平衡上加以考虑，质量控制也只要在总体上加以检验。

2. 试题的统计分析

语言测试的统计分析，是随着测试标准化改革而发展起来的。统计分析的目的是从测试的结果中取得关于测试本身质量的信息。统计分析的项目有如下一些：

1）答对率、难易度和区分度。答对率也称通过率，指答对的人数与全体受试者之比，其值为0～1。假如一道试题的答对率是0.5，则意味着有一半人答对，一半人答错。难易度是另一个数据，可用公式从答对率推算出来：难易度可用希腊字母Δ来表示，Δ的值为1～25，中间值为13，数值越小表示试题越容易；Δ低到9时，表示有84%的人答对；Δ高到17时，表示只有16%的答对。对试题难易的评估，可以直接根据它的难易度。区分度常用的计算公式是结合本题的答题情况，算出双列相关系数，用$r_{bis}$表示。$r_{bis}$的值为－1～1。如果一道试题的$r_{bis}$是1，意味着高档次的受试者都答对了本题；而如果$r_{bis}$是－1的话，则意味着低档次的都答对了，高档次的反而答错了。这两种极端情况只是理论上具有可能性，实际上一道试题的$r_{bis}$达到0.3就符合要求，超过0.4就说明这道试题的区分能力很不错了，$r_{bis}$是负数的话，这道试题是不能用的。对试题进行微观分析（即分析每一具体的题目）也好，对试题进行宏观分析（即进行全卷或部分大题的分析）也好，答对率、难易度和区分度都是统计分析的指标。例如全卷的答对率一般要求是0.5；选择题由于有猜测的原因，答对率指标应稍高一点，达到0.6左右；区分度总的要

求不能低于 0.3，而且 $r_{bis}<0.3$ 的试题要求控制在总题的 5%以下。

2）信度。信度是全卷或大题统计分析的又一个重要质量指标。信度可用我们前面介绍过的公式计算，即：

$$r = VAR(t)/VAR(X) \text{ 或 } 1 - VAR(e)/VAR(X)$$

当信度 $r$ 的值为 0.50 或 0.75 时，误差过大，一般的测量均要求信度在 0.90 以上的值。这个值与试题量、标准差、答对率都有关系，试题数越多，标准差越大，信度就越高。

3）相关矩阵。可以把一些相关系数算出并排成矩阵，这也是语言测试后经常使用的一种统计分析方法。例如我们可以用矩阵方法进行因素分析，这是测试效度的一个方法，例如假设某次测试共考两个因素：知识题（K 题）考的是因素 1，运用题（U 题）考的是因素 2，然后将试题与因素列成矩阵的纵坐标和横坐标：

| | 因素 1 | 因素 2 |
|---|---|---|
| K1（语法填空） | 744 | 256 |
| K2（选择题） | 543 | 457 |
| U1（阅读理解） | 404 | 596 |
| U2（听力） | 199 | 801 |
| …………………… | | |
| 因素负荷：总 40.0% | 18.7% | 21.3% |

通过矩阵表了解到，18.7%的分量测试了语言知识，21.3%的分量测试了语言运用能力，都是属于有效因素（低于 10%为无效因素），而且因素 2 大于因素 1，有一定的效度。但是因素总负荷（有效因素的负荷总和）太低，因此这份试卷是不理想的。

4）适宜性分析。这种分析利用概率原理，在人的能力和试题的难度的关系中，算出每个受试者的能力及每道试题的难度，然后推论每个已知能力是多大的受试者，答每道已知难度有多大的题，答对的概率是多少，从而得出全体受试者对全套试题的答对概率模式。如果整个受试群中有几个受试者的表现不符合这个概率模式，出现超出限度的误差，我们说这几个人不适宜这套试题；如果整套试题中有几道试题不符合这个概率模式，出现超出限度的误差，我们说这几道试题不适宜于这个受试群。不适宜的受试人数和试题数，占全体受试者及全卷试题数的 5%左右，这是允许的；接近甚至超

过10%，这个测试的效度、信度就很有问题了。

## 第四节 语言测试中不同题型的具体命题问题

### 一、试题和题型

试题和题型是不同的概念：试题是指测试什么，怎么测试，题型则单纯指怎么测试。测试同一种目标因素的试题可以用不同的题型；同样的题型也可测试不同目标因素的试题。从试题的分类来说，大约有A、B、C三类，A类试题既是离散考点的，又是单技能的；B类是综合考点的，但是单技能的；C类既是综合考点的，又是综合技能的。而语言测试中的题型种类却难以列举，因为命题人随时可以创造新的题型。常见的题型有选择题、是非题、排序题、改错题、填空题、听力题、改写题、写作题等等。

一套测试的信度、效度，要靠试题去落实；一套试题的信度、效度，要靠题型去落实；而题型的信度、效度，则要靠命题去落实。所以命题工作的努力方向，不仅仅是把试题按照它所选定的题型去保持其原来应有的信度和效度（尽管这已经是很不容易了），而是进一步探索如何突破既定题型，超过既定题型应有的信度和效度：对客观题，着力提高效度；对主观题，着力提高信度。下面讨论的是一些具体题型的命题步骤、命题要求等问题。

### 二、MC类题的命题

#### 1. MC类题的组成

标准的MC类题由两部分组成：题干(stem)和选择项(options)，选择项中有答案项(key)和干扰项(distracyors)。一个问题，往往提供4个可能的答案，问题就是题干，4个可能的答案就是选择项，其中正确答案是答案项，另外3个是干扰项。虽然理论上认为每道题的选择项越多越有效度，由于命题的困难，一般正式的测试用4个选择项被认为是恰当的。一般规定，每题中有一个是答案项。有的题可能有多个答案项，也有可能没有答案项，这种不定选项数的MC题比定死为只选一个的MC题要活得多，理论上也更有效度，但是评分比较困难，一般很少用。

2. MC 类题的命题

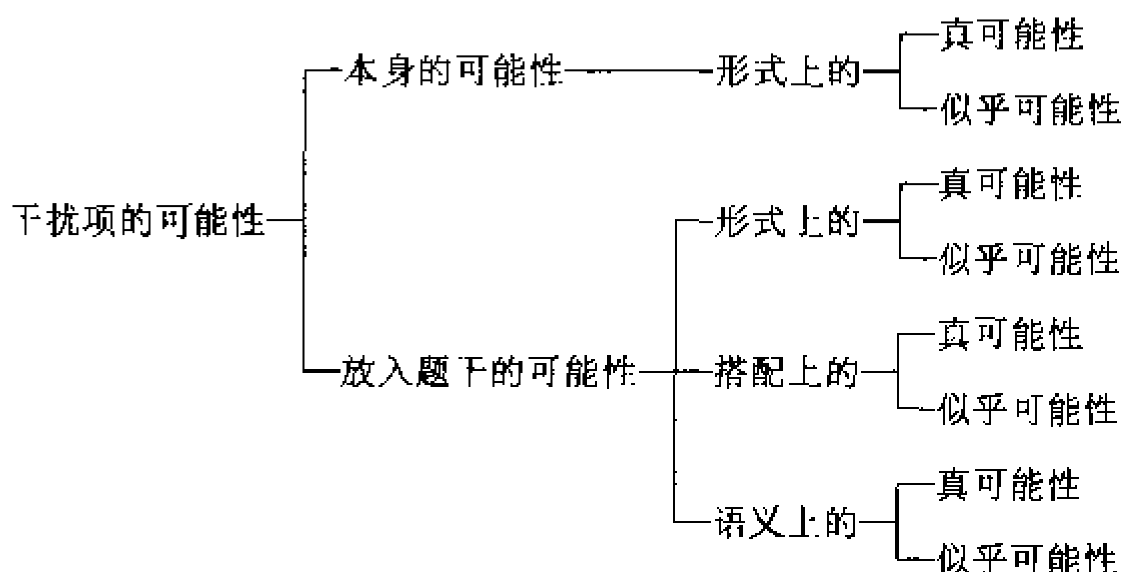
MC 类题的命题一般采用定考点、制定题干和答案项、制定干扰项、排列选择项等步骤。

1）定考点。考点总的分布要根据总的设计，具体到给每道测试题定考点，也应在总体分布的大前提下考虑。例如语法题的考点，首先要以目标语语法体系中应考范围为根据，这个范围，应在测试大纲和教学大纲中得到确定，可以将这些语法考点以表格形式列出；其次，可以以测试对象受试群常犯的语法错误所反映出来的语法薄弱点作为定考点的根据。从这两个根据出发，就比较容易定出本次测试的语法考点了。决定具体的每一道题的考点时，要考虑到 3 个问题，即 ① 考点是否属于测试范围内？也就是说，考点必须明确；② 考点是否只有一个？一般来说，MC 类题每题只考一个考点，或顶多两个能互相增补的考点，所以命题时不能把一些额外的考点牵涉进去，更不能把假考点和真考点都用在试题上；③ 考点是否值得考？命题人应该了解到在测试范围内的所有考点中，哪些值得考，哪些不值得考，如果要精简的话，命题人应该知道精简哪些考点。

2）制作题干和答案项。题干加答案项，便是考点的体现。一般来说，对填空题只要写出完整、正确的句子，然后按考点的要求把需要抽出的部分抽出作为答案项就行了；对改错题来说，也应先把正确的句子写下来，再按考点的要求将对的改成错的，这部分就是答案项。要注意的是，每题必须有一个答案，每题也只能有一个答案，除非是无答案和多答案的 MC 类题。题干和答案项的语言运用要规范、简洁、得体，语言运用不正确，往往会造成无答案或多答案的后果。另外，出题时要避免偏颇性，如果试题牵涉到的文化背景或题材领域，使一部分受试者感到比较熟悉，而使另一部分受试者感到比较生疏，这就会造成事实上的不公平。

3）制作干扰项。决定了考点，制定了题干和答案项之后，接着就是写干扰项。有一种错误的认识，以为干扰项只不过是错误的答案，只要随便地把正确答案改成错误答案就行了。其实制定干扰项可以说是 MC 类命题工作中最难的一个环节了，因为制定干扰项要以受试者围绕本题的考点可能会产生的反应作为根据。而要知道受试者的反应，要借助多种理论，包括目标语本身的体系，目标语与母语的对比，第二语言的学习，测试心理学等方面的理论，也要借助实际经验，包括目标语的教学经验，以往类似测试的操作经验等等。制作完了干扰项，命题可以说实际上已经完成。每一个干

扰项都应该具有似乎可能性(plausibility),似乎可能性是看起来的可能性,可以是真的,也可以是假的。就是说,似乎可能性可以以真为可能性作为它的前提,也可以没有这个前提。干扰项的似乎可能性是两个层次 3 个方面的:两个层次是本身的可能与放入题干的可能;3 个方面是形式方面、搭配方面和语义方面的可能。干扰项不可以两个层次 3 个方面都是可能的,这样就变成了答案项了,但不排除某一个层次某一个方面的可能性。干扰项的似乎可能性可用下表表示:




似乎可能性是一个程度问题:干扰项的似乎可能性太低,就失去作用;干扰项的似乎可能性太高,却又有成为答案的危险。例如下面这道题的干扰项的制作注意到了各个方面的干扰因素:

| | 选择项 | A | B | C | D |
|---|---|---|---|---|---|
| 与题干是否相容 | 语法上 | + | + | + | − |
| | 搭配上 | + | − | + | + |
| | 语义上 | − | + | + | − |

制作干扰项要利用受试者可能犯的答题错误,利用的渠道有 3 点:① 利用目标语学习者的典型弱点,学习者犯过的错误或可能犯的错误,常常是制作干扰项的依据;② 利用对词语的误解或错误联想,这种误解往往是由于抓住只字片言,望文生义而引起的,与学习过程中的典型弱点没有关系;③ 利用近似词语之间的混淆,这种近似,包括词性、语义或者结构上的近似。

4) 排列选择项。每道试题的选择项先后排列,是有讲究的,否则会无意中为猜测答案提供了暗示。一般选择项的排列要考虑到以下几个方面:① 选择项的制作最好要求长度一致、类属一致、结构一致、难度一致。如果

选择项中是两长两短，可以按照"短短长长"、"长长短短"、"短长短长"、"长短长短"的方式排列，目的是让受试者明显地看出来，选择项按长短排列，跟意思或答案毫无关系。② 按数字顺序排列。如果选择项和数字有关，这四个选择项可按数字的大小，或者从大到小，或者从小到大。时间、日期、年代等等作选择项时，也可按这种规律排列，不让受试者从选择项排列中得到某种暗示。③ 把选择项在形式、意义、结构上有相似、相配或对立点的放在一起，以突出这种相似、相配与对立，使选择项尽可能地平衡、整齐，有利于受试者公平地作出选择。④ 在难易程度上作好安排。一般来说，较易的答案项宜放在后面，较难的答案项宜放在前面，可以让受试者看了 4 个选择项之后才能作答。较易的如果在前面，受试者只看了一个就行了。不过这种方式也只能适可而止，不能形成规律。

总之，MC 类题命题的公式化，正是我们当前的努力方向：要把人从一种工作中解放出来，并且使这项工作做得更好，就必须探索这种工作的规律，把这种工作的规律数码化，亦即方程式化、公式化。数码化其实是一切科学发展的最终归宿，语言测试也不例外。在语言测试命题领域中，现阶段我们所能做的，还只限于探索语言测试中比较有规律的部分的规律，使这些规律数码化。

## 三、阅读理解题的命题

### 1. 阅读理解能力

阅读理解能力是一种领会性能力(receptive ability)，不是产出性能力(productive ability)。对产出性能力的测试，可以通过写作，对言语活动的结果进行评价，是可以直接观察到的；而领会性能力是无法直接观察的，阅读观察是受试者内心的心理活动。在这个心理活动的过程中，受试者的大脑在想些什么，利用了哪些知识，如何进行推理和分析等，至今都未找到令人满意的答案。

### 2. 阅读理解题的测试内容

阅读理解能力是一种领会性能力，虽然无法直接观察到，但从现在的认识看，至少包括以下一些能力：① 认识词、短语，在语境中理解词和短语的意义；② 理解句子结构；③ 掌握所读材料的主旨和大意，了解说明主旨和大意的事实和细节；④ 根据字面意思，根据所读材料进行一定的判断和推论；⑤ 在理解个别句子意义的基础上，理解上下文的逻辑关系，理解文章

的深层含义。上述能力的表示，似乎很清楚，但实际上阅读理解能力仍然是一个难以说清的东西：第一，上述③、④、⑤能力的具体过程仍然是不清楚的；第二，阅读过程肯定要比上述 5 条要复杂得多。阅读不仅仅是一个语言问题，与一个人的普通常识有关，与一个人的思维方法有关，与一个人的文化背景也有关。所以现在的大部分阅读测试是基于这样的认识：如果你没有看懂某一点，你就答不上某个问题。至于你是如何看懂的，组织测试的人并不完全清楚。尽管这样，实际操作时，在测试内容规范中将上述能力进一步细化为下列 3 类 9 项技能：

(1) 辨别和理解中心思想和重要细节

1) 理解明确表达的概念或细节；

2) 理解隐含表达的概念或细节（如下结论、作判断、进行推论等），通过判断句子的交际功能（如请求、拒绝、命令等）来理解文章；

3) 辨别文章的中心思想（如找出能归纳文章中心思想的要求）；

4) 理解作者的观点和态度。

(2) 运用有助于理解的语言技能来理解文章

1) 猜测词和短语的意思；

2) 句层次的理解（如理解句子所表达的原因、结果、目的、比较等）；

3) 篇章层次的理解（如运用词汇的、语法的承接手段来理解文章各部分的关系）。

(3) 运用专门的阅读技能理解文章

1) 通过略读了解文章大意；

2) 通过查阅寻找某一信息。

从测试内容规范可以看出，阅读理解题主要测试对阅读内容的理解，命题时考点的确定应体现理解的不同层次，不但要理解字面的意义，而且要理解字里行间隐含的意义，理解作者的观点和态度等等。由于阅读理解测试是一种间接测试，阅读理解测试的效度受到两个因素的制约，即阅读材料的是否适当和命题的质量高低。艰深的文章，如果题目出得很容易，即使受试者全部答对，也不能由此推断出受试者已经具备了理解同类文章的能力。该测试的不测试，不该测试的测试了，不能说测试是有效度的。

**3. 阅读理解题材料的选择**

选择阅读理解题的阅读材料首先要考虑的是材料的合适性。所谓合适，是指内容和语言表达两个方面都要合适。内容是否合适主要看所选的

材料是否符合教学大纲,是否符合测试的目的。由于阅读理解测试中所用的阅读材料只是受试者将来所需要阅读的材料的一个抽样,因此阅读理解测试中所用的材料要具有真实性,包括交际目的的真实性,文体的真实性和语言运用的真实性。第二语言教学的最终目的是使学习者能够用目标语作为工具通过阅读获取本专业所需的知识和信息,这样,阅读材料一般不选择小说、诗歌类体裁的文章,而以学术性文章、社会性文章为主。在这一前提下,选材应包括议论文、叙述文、说明文等,以体现多方面的阅读需要。除此之外,阅读选材还要考虑到以下几点:① 选材应有利于测试各个层次的阅读能力,因此最好选择有新意的、观点上有转折的材料;② 要选择有一定信息密度的材料,而不选平铺直叙的材料,以便从不同的角度进行提问;③ 选出的材料避免涉及过深的专业知识,避免过于偏颇的文化背景,以保证对不同受试者的机会均等。

从语言表达来说,要有适当的难易度。这一点的掌握是不容易的,对语言程度的高低,主要还是凭命题人的主观判断,即使有词汇大纲、语法大纲等,事实上还是缺乏科学的根据。阅读材料中的语言也应具有真实性的特点。这种真实,是指对特定的情景而言的准确和得体的语言。真实性包含了规范性,这种规范性不是架空的绝对规范性,而是具有社会语言学定义的规范性。真实性并不意味着材料中的语言不能加工改写,事实上中低级语言测试中的阅读材料,很多都不可避免地要进行改写。不过,改写过的语言必须自然而不露痕迹,能保留所属篇章类型的特征。

总之,阅读材料的选材原则可以归纳为以下 3 点:

1) 题材广泛,可以包括人物传记、社会、文化、日常生活、科普常识等等,但是所涉及的背景知识应能为受试者所理解;

2) 体裁多样,可以包括叙述文、说明文、议论文等等;

3) 语言难度和词汇量符合教学大纲规定的范围。

**4. 阅读试题的制作**

阅读理解一般也是用 MC 类题的方式进行测试,与语法试题等不同的是,阅读理解题的焦点是直接放在信息、意义上的,应该从信息获取的角度去考虑和解释,所以,阅读理解中的 MC 试题干扰项的设计,也应以受试者从材料中获取信息的过程中可能会出现的差错为依据。既然阅读理解测试的是信息的获得,那么阅读理解题的考点应该以信息为目标,要以未知信息为考点,也就是说,不要问受试者本来就知道答案的问题。在阅读材料本身

含有多种层次多种范畴的基础上，要尽量在试题中体现出多种层次多种范畴的信息。另外，在MC干扰项的设计上，考点对材料内容的覆盖应尽可能全面，不能明显地只集中考材料的一部分而置另一部分于不顾。

从出题技巧上说，要充分考虑人在语言交际中接收信息时会理解错误的原因。得出错误信息可以从3个层面上找原因：① 把具体的词语弄错了；② 词语没弄错，但词语的意思弄错了；③ 具体词语没弄错，一词一语的意思也没弄错，但却把某点意思或信息在语篇中的意义弄错了。阅读理解试题中产生信息理解错误主要是第三个原因造成的。这样，在出题时就可以从以下几个方面入手：① 对信息的错误联想，阅读理解试题引发联想的是意义，某点意思是阅读材料里本来有的，却把它和材料语篇里没有的、不相干的一些别的意思联系在一起，产生错误联想，造成语篇层面上的信息理解错误；② 对信息的错误判断，在阅读理解题里体现为真假混淆，利用的是材料里的信息的似是而非或似非而是，特别是在考潜信息的情况下，这种技巧用得很多；③ 弄错信息间的关系，信息与信息之间，以一定的关系互相组合，如主次关系、因果关系、目的和手段的关系、前提和结论的关系等等，这些关系，也是篇章含有的信息的一部分，可以利用信息之间的颠倒的关系来设计干扰项；④ 推理或演算错误，给受试者前提，再给受试者结论，让受试者判断，根据推理或演算这结论是对还是错，前提在材料里给，可能的结论以选择项的形式出现，干扰项建立在错误的推理或演算上。

总之，阅读理解中的MC类题，是不同于前面一节中的MC题的：前者以信息为焦点，后者以语码为焦点；前者不是测试语言形式结构的试题，而是测试信息意义的试题，后者主要测试语言的形式结构，是内容和形式的较好结合；前者可归入直接测试的范围，后者只能说是一种间接测试。阅读理解测试虽然可以采取MC题型，但却不能出成语法、词汇或句子的释义题。

### 四、写作题的命题

#### 1. 写作试题的特点

写作试题是一种测试语言的产生性交际功能的试题。语言测试中的写作试题，不能只测试语言能力，而要测试交际能力，这个交际能力包括语言能力，还包括语篇能力、语用能力，所以写作试题测试的是受试者多方面的知识、技能和能力。同时，语言测试中的写作试题，测试的对象毕竟是语言，交际能力尽管也包括写作技巧，但修辞、文章方面的技巧要截然分开。

写作可以看作是现代语言测试可能取得最高效度的试题。因此，作为写作试题的特点来说，也是很明显的，就是效度高，信度低。

(1) 效度高

写作试题的高效度体现在下面几点：

1) 现代语言测试科学理论认为，听说读写的试题都是直接试题，具有较高的效度。其中写作和说话的试题称作产生性运用的试题，而阅读和听力则被称作接收性运用的试题，前者比后者的效度更高：测试接收性能力，未必能测试到产生性能力；而测试产生性能力，则一定能测试到接收性能力。读懂的不一定能写出来；能写出来的则必定能读懂。因此，测试写作能力的同时等于在测试阅读能力。

2) 现代语言测试科学理论认为，语言测试考点具有4个效度层次和3个因素范畴：4个效度层次从低到高依次是：单词、词组(短语)、句子和语篇，3个考点因素从低范畴到高范畴依次是语法、搭配和意义。考点测试到哪个层次，必然同时测试到其以下的所有层次；考点以哪一个因素为焦点，也必然能同时测试到比它范畴低的因素。写作试题中，只要测试的是语篇层次的写作，就自然把句子层次、词组层次、单词层次全部测试到；只要测试的是以意义即信息为焦点的写作，就自然也测试到了搭配、语法。从这个意义上说，写作试题能同时测试语言的各个层次和各个因素。

3) 现代语言测试科学理论认为，语言表现评估的标准是准确性、流利性和得体性。准确性是指语言符合语法规则，即符合语言形式系统自身的规律，这属于交际能力中的语言能力；流利性是指语言运用过程中的顺利程度，即运用中的效率，这不是语码系统本身的向题，而是语言在运用中的属性，牵涉到的是语篇能力和语用能力；得体性是指语言的运用是否符合语境情景的要求，这就直接属于语篇能力和语用能力的范畴了。写作试题可以同时测试到语言的准确性、流利性和得体性。

(2) 信度低

写作试题的最致命的弱点是信度低。这种试题属于主观题，评分无法客观，信度也自然无法得到保证。因此，实际测试中，应根据条件选择现有办法中最能保证信度的一种，克服评分中的3个不一致。

从理论上说，接收和产生，是语言的两个根本方面，缺任何一方都是不可以的：语言测试只测试接收不测试产生，说到底不能算是语言测试，没有写或说的试题的语言测试，本身的效度是要大打折扣的。目前的语言测试，

几乎都有测试产生性能力的试题如写作等，最主要的理由是这种试题有极好的反拨效度(反拨效度指的是考试对教学是不是有良好的反拨效应)。在当前有广泛可行性的所有各种试题中，这种试题是惟一能带给教学纯粹正面而不带负面的、最全面的反拨效应的试题，因而也是任何其他试题所不能取代的试题。

2. 写作试题的命题原则

写作试题既然是高效度试题，就应该尽可能地保证其最高效度，发挥其最大的效度优势，这是写作试题命题原则的最根本的一点。除此之外，还要注意以下几点：

1) 试题以信息表达为测试目的。写作试题与改写、串句成文等练习是不同的：前者从意义出发，以特定意义的表达为测试目的，测试的焦点是信息；后者从形式出发，以特定的语言形式(如词汇、句型和语法)为测试的目的，测试的焦点是语法和词汇。而意义的测试，又同时可以达到测试语言形式的目的。

2) 试题中出现的情景，对受试群来说应该是真实的。所谓真实的情景，是指恰当的(relevant)情景，是受试群现在或将来很有可能会遇到的情景。情景的真实还意味着：① 现在的真实情景不同于阅读的真实情景，一个人通过阅读获取信息的渠道或情景，正在不断增多，而需要通过现在来获取信息的渠道或情景，仍然还是有限的几种；② 第二语言的真实情景与母语的真实情景是不同的，特别在第二语言测试过程中，仍要注意两者的差异。例如测试受试者用母语写作小说，那是真实的；如果要求受试者用第二语言去写作小说，这种情景就显然是不真实的了。

3) 测试的层次应该是语篇层次。写作试题应该在语篇层次上测试信息的表达。例如看图写句子和看图写文章是完全不同的两码事：前者虽然也产生单词和句子，但单词和句子的产生却没有语境，这种试题的测试只能说是达到单词或句子的层次；后者的图中需要提供更多的内容，语篇的定义是语境和情景中的语言，是既具有平面的维度，也具有立体的维度的语言。测试的层次达到了语篇，试题才能说是真正地测试了产生性的交际运用。

4) 试题中体现的应该是互动性的行为。交际是信息的传递，信息的传递不可能是单方面而必须是双方的，通过字里行间的意义磋商(negotiating for meaning)才能获取信息。写作试题需要设计的是互动性的写作行为，例如写作的试题是对来信的答复，那就是体现了互动性；如果只是个人的写作，也应

该明确知道是写给谁看的,即让受试者有明显的对象意识去进行写作。

5) 语言的制约应来自交际情景。写作试题对语言的控制,自然地来自交际情景,而不是来自人为的对语言形式的主观选定。写作试题应体现下列模式:

交际情景——(制约)——要表达的意义——(制约)——要用的语言形式

先有情景,情景决定要表达的意思,要表达的意思再决定要说的话的形式。这个模式,是真实生活里产生性地运用语言的模式。

3. 写作试题的设计

写作试题由以下几个部分组成,每一部分的具体设计组成这个写作试题的总设计。

1) 情景。写作试题的设计中情景的设计是关键的第一步。写作题的情景设计,除了依靠测试大纲、教学大纲等,依靠命题人对受试群的现状和将来的需要的了解以外,还必须依靠命题人的想像力。在情景必须真实这个大前提下,命题人的想像力是非常重要的:想像力的作用,在于想像出种种有新意的,但是仍然是真实的情景。写作题对想像力的要求,对任何其他类型的试题的设计,都是不需要的。

写作试题中的情景可以分为无需假设情景和需要假设情景两种。传统的只出一个题目让受试者去写的方式,实际上向受试者提供了一个假设的情景:尽管受试者是以自己真实的身份写文章,但让他写文章,而不知道写给谁看,或者只知道写给评卷人看,这就不是实际生活中的交际情景了。因此,写作试题中常常需要假设情景,这可以通过不同的手段,如语言、图画、录像、多媒体技术等等,其中以前两种为主。用语言说明情景,要告诉受试者以什么身份,向以明确身份存在的对方写一点什么。说明要清楚、简洁,设计的情景与人物关系不能太复杂,能让受试者明白试题要求他进入的是什么情景。用图画提供情景,要比语言说明更加形象化,但图画也往往需要用语言作进一步说明,以便使情景更加明确。

2) 馈入。馈入是指受试者在写作之前,给他看的任何形式的材料,亦即给受试者提供他要输出的内容。如果所提供的情景能充分推断输出的内容,则馈入无须另加。馈入有两方面的作用:第一,对输出的内容具体化,有时虽然提供了情景,但是输出要求的范围还是太宽太泛,这时需要提供一定的馈入;第二,对受试者的输出加以必要的限制,可以使受试者的写作内容

有一定的统一性和可比性,能保证评分达到要求的信度。

馈入的设计要遵循两条原则。第一,馈入限于内容,不提供表达的形式。如果受试者从试题的馈入中直接取得表达形式,试题的效度就要受到影响。语言形式应该让受试者自己根据情景自由选择决定。第二,即使是内容的提供,也要适可而止,能起到上述两个作用就可以了,就是说,在内容选择上面,应尽量在现实容许的范围内,给受试者尽量多的自由。所以,试题的馈入不应是现成的形式,也不应是现成的内容,而要让受试者经过一个领会和重新处理的过程,才能得到他具体可以输出的信息,然后编码输出。

3) 输出。情景和馈入的设计,是为了取得成功的输出,写作试题的成功与否,最终落实到输出上。成功输出的标准是:输出应该是在真实情景下,语篇层次上,互动性的信息表达。选择试题的输出实际上是一种抽样,是测试受试者的能力的表征行为的抽样。作为抽样,要符合下列要求:① 抽样必须有一定的量才有效,在可行性允许的情况下,应尽量使受试者的输出量多一些,例如大规模的作文考试,常常有字数的规定; ② 抽样尽量体现多样化的要求,每次测试中能测试到的抽样种类多一些,效度就高一点,例如目前国内高考的语文试题中,除了大作文之外,往往还有小作文,目的是追求抽样的多样化。

科技的发展,特别是认知科学、教育科学、人工智能研究、信息技术等方面的发展,给语言教学和语言测试带来了观念上的影响,这种影响可以归纳为三方面:

1) 语言教学和语言测试的个体化。当我们力图把第二语言教学的中心从教师转移到学习者以后,就会进一步发展为个人化的学习(individualized learning),而个人化的学习的实现则要求做到以下几点: ① 拥有学习的控制权。每个学习者可以灵活地决定自己的学习时间、学习进度、学习顺序、学习方式、学习策略、学习分量,乃至参与课程计划和教材的编辑。② 具有适应性的课程。要将课程设计按不同水平、不同类型、不同渠道作不同顺序的排列,构成不同路线、不同焦点的课程,以适应学习者个人不同的学习需要、学习风格、学习兴趣和学习能力。③ 能获得个人化的反馈。学习者可以经常获得关于自己学习的信息,如学习诊断报告、学习者图像描绘、学习进程路线图、成绩报告以及针对个人的学习指导、学习材料和学习方法的介绍、以往学习的总结、以后学习的处方性意见等等。

2）语言学习和语言测试的真实性。应该根据现代科学对语言学习和语言运用的新认识，去设计语言学习和语言测试，尽可能地接近学语言、用语言的真实心理状态。要实现这种真实性，也应该做到以下几点：① 对情境化的要求越来越高。情境化不仅是教学和测试的方法问题，而且是教学和测试的性质问题。随着各种学科的发展，尤其是信息技术的突飞猛进，语言教学和语言测试的情境化会获得新的维度和新的高度。② 对互动化的要求越来越高。互动化不仅存在于教师与学习者之间，还存在于学习者与学习者之间、说话者与听话者之间、学习者和课堂以外的世界之间、学习者新的经验和旧的经验之间、新的知识和旧的知识之间。实现语言教学和语言测试的互动化，更符合学习者学习语言和运用语言的心理活动。③ 对智能化的要求越来越高。认知心理学的发展，使人们进一步认识到学习第二语言的心理过程以及智力在这个过程中的作用。因此，第二语言的教学中应树立把学习者看成是有智慧的人的教学风气；在第二语言的测试中应抛弃那种出题人无须用脑，答题人也无须用脑，把人越考越笨的考试方式。

3）语言学习和语言测试的过程化。由于近年来种种相关学科的发展，使人们对在第二语言学习和运用时的心理活动更感兴趣，更加了解。对第二语言的学习者来说，掌握学习和运用第二语言的活的过程，比直接掌握死的"成品"更为重要。语言的学习是以过程为焦点的学习，本质上就是要获取在过程中运用的能力，而非获取拿在手中的现成的知识。过程为焦点的语言教学和语言测试，会带来以下的变化：① 学习本身将会成为一种无所谓终点成果的过程，第二语言的学习，会成为个人有需要时随时可以拾起来，又随时可以放下去的持续终生教育（on-going education）的一部分。② 学习和测试将合成一体，测试会被连续性的形成性评估（continuous formative assessment）所替代，不以测定某一时刻的水平为目的，而以测试动态的进步为宗旨。测试反馈的不应是静态的位置（static position），而应是动态的轨迹（dynamic trajectory）。

科技发展会给语言教学和语言测试在应用上带来方法、手段上的革命。电脑应用、多媒体技术、人工智能应用、自然语言处理等学科的发展，为语言教学和语言测试开辟了多方面的实践途径。

1）测试前期的应用。今后的发展一是使题库具有灵活性，一是建立资料库以取代题库，例如建立测试部件库、作业库、适应不同目的的材料库等等。随着语料库语言学的发展，资料库的建设越来越现实。与此同时，新科

技的发展,将改变目前占大部分的非调适性的测试,试题结构中的调适性设计问题可以解决,且调适的程度会越来越高,调适的范围会越来越大。

2)测试中的应用。电脑在测试过程中可以发挥5种功能:① 感知(sensing)。电脑的感知主要通过输入手段,而电脑的输入手段正在日新月异地变化:键盘、下拉菜单、图标、鼠标、触摸屏、扫描仪、声音识别仪等等。② 记忆(remembering)。电脑的记忆容量越来越大,记忆手段越来越多。CD-ROM、WORM 光盘、影视盘等使电脑的储存容量以几何级数增长。③ 决定(deciding)。电脑作出决定,以对输入及记忆中的信息的数学即逻辑运算为依据。微处理器的能力正在向更高层次、更高速度的方向发展。④ 行动(acting)。电脑的行动即电脑的输出主要指控制显示(包括文字、图像、声音的显示),手段越来越丰富。⑤ 传送(communicating)。电脑的联网使远程教学成为可能。上述5种功能,会使语言测试的实施发生彻底的变革。

3)测试后期的应用。测试后期的工作主要指评定分数、统计分析分数、解释分数。这3项工作可以依靠电脑的上述五种功能。

总之,由于新科技的影响,语言教学和语言测试将会不断地向前发展。最终,语言测试将不再服务于分数,不再服务于官僚,不再服务于考试费,而只服务于教育。

## 参考书目


1. 桂诗春.心理语言学.上海外语教育出版社,1985
2. 桂诗春.应用语言学.湖南教育出版社,1988
3. 桂诗春.实验心理语言学纲要.湖南教育出版社,1991
4. 桂诗春、宁春岩.语言学方法论.外语教学与研究出版社,1997
5. 王初明.应用心理语言学.湖南教育出版社,1990
6. 冯志伟.数学与语言.湖南教育出版社,1991
7. 李筱菊.语言测试科学与艺术.湖南教育出版社,1997
8. 刘润清.语言测试和它的方法.外语教学与研究出版社,1991
9. 杨惠中,C. Weir.大学英语四、六级考试效度研究.上海外语教育出版社,1998
10. 〔英〕阿里·戴维斯著.语言测试原理.任福昌等译.经济科学出版社,1998

# 第四章　中文信息处理

## 第一节　中文信息处理概述

### 一、中文信息处理及其研究内容

1. 什么是中文信息处理

语言的最重要的社会功能是作为人与人之间相互交流的工具，但随着计算机技术的发展以及计算机在各个领域中运用范围的扩大，语言这个工具的使用范围也逐渐扩展到用于人与计算机之间的信息交流，因此如何让计算机像人一样理解自然语言的问题日益受到人们的重视，并逐步形成了一门跨语言学、计算机科学等许多学科部门、运用计算机技术来研究和处理自然语言的交叉性学科——自然语言处理(natural language processing，简称 NLP)。自然语言信息处理的主要目标是要让计算机懂得和理解人类的自然语言，也就是要使计算机能模拟人的语言机制。50 多年来，这项研究取得了长足的进展，成为一门重要的新兴边缘学科。它的出现"使得语言学在现代科学体系中的地位有了明显的变化，使语言学由一门基础科学变成了领先科学，获得了与数学、哲学同等的地位，语言学将成为人文科学发展的突破点和生长点，它的重要意义已经为越来越多的人所认识"①。

中文特指汉民族的语言文字，也就是汉语的书面形式和口头形式，是众多自然语言中的一种。作为自然语言处理研究的一个部分，中文信息处理指的是用计算机对中文(包括汉语的书面语和口语)进行存贮、转换、分析、传输等加工以形成各种信息处理系统的科学，是自然语言信息处理这个应

① 冯志伟. 自然语言的计算机处理. 上海外语教育出版社，1996. 3

用语言学的分支在汉语汉字应用研究中的体现。随着科学技术的发展，中文信息处理技术已深入到社会生活的各个方面。中文的书面形式是汉字，由于中文在语音、词汇、语法等方面有它的特殊性，在用计算机来处理中文记载的各种信息的时候，就自然会提出特殊的新问题，对这些问题的研究，构成了中文信息处理的基本内容。

中文信息处理是一个多层次的结构，主要包括中文“字”处理、“词”处理和“句”处理几个方面。以书面汉语来说，汉字是中文最基本的层次，汉字信息处理主要包括汉字的编码、输入、输出、汉字的自动识别、汉字的统计特性研究等，是中文信息处理的基础性工作和关键环节。汉语信息处理是汉字信息处理的进一步发展，它要在汉字输入的基础上，研究汉语的词汇、句法、语义、语境的自动处理问题，是中文信息处理的高级阶段。这一阶段要研究汉语词的自动切分和处理，汉语短语的自动处理，汉语句子的自动分析和生成，汉语语义的自动分析和加工，以至于汉语篇章的自动处理等。中文信息处理需要以大量的语言知识、百科知识为支撑，让计算机对中文信息的人脑处理过程进行模拟。对于口头形式的汉语，中文信息处理还要研究语音的自动合成与识别、自动文语转换等问题。

2. 汉字信息处理

汉字信息处理简单地说就是让计算机能接受并处理汉字，它是中文信息处理的关键和需要突破的核心技术，主要包括 3 个方面的内容：汉字输入、汉字存储和汉字输出。

(1) 汉字存储和输出

存储到计算机里的任何信息都是用“二进制数”来表示的，即只用 0 和 1 这两个阿拉伯数字来表示所有的数字。二进制数的每一位在计算机里叫做一个比特(bit)，8 个比特为一个字节(byte)，计算机存储器的容量就是以字节为单位的。世界上第一台计算机 1946 年诞生于美国宾夕法尼亚大学，其软件和硬件都是针对处理英语而设计的，记录英语的拉丁字母(包括大小写)总共才 52 个，再加上各种常用的符号，总共不过 100 多个，因此存储英文字符以及常用符号，只需 7 位二进制数就可以了。而记录汉语的汉字数量大，通用汉字有 7 000 多个，另外还有各种各样的符号。由于数量巨大，存储汉字得用两个字节，即用 16 位二进制数作为一个汉字的存储空间，用两个字节表示一个汉字，比起英文只用一个字节来表示一个英文字母要困难得多。

信息在计算机内是以代码形式存在的，对信息的处理就体现为对代码的处理。目前在计算机中最普遍采用的西文字符代码是 ASCII 码，计算机处理汉字时，也必须先将汉字代码化，然后对汉字代码进行处理。为了能在计算机中进行汉字处理，首先得对汉字进行编码。汉字编码可分为内码和外码两种。

汉字内码又称为汉字机内码，指在计算机内部进行存储、传递和运算所使用的汉字编码。目前使用的汉字机内码体系是国家标准 GB2312，共收录汉字 6 763 个，机内码长两个字节，共 16 位二进制位。计算机里的汉字是占两个字节的二进制数所表示的汉字字形，这种字形称为"数字化字模"。同时在计算机的字库里每个汉字字形都用数字化的点阵方式存储着。

汉字输入计算机以后，要先转换成由相应的计算机系统所确定的机内码，然后再还原成汉字进行输出。汉字的计算机输出有两种形式，一种是屏幕显示输出，一种是打印输出。当要输出汉字时，数字化字模就转变成字库里的数字化点阵，并在屏幕上将汉字显示出来。为了使输出的汉字规范和统一，达到汉字字体的正确性、整幅汉字的一致性、实际字形的清晰性、笔形部件的规范性和使用效果的美观性，有关部门研制和规定了统一的汉字字形(包括宋体、仿宋体、楷体、黑体等)的点阵标准，并统一提升为国家标准。"点阵字库"有很多优点，突出的优点是制作方便，输出速度快；但也有缺点，那就是不能随便放大缩小。为解决这个问题，经研究发现，汉字的每一笔笔画，其周围的轮廓，都可以看作是数学上的一个矢量。数学上所谓一个"矢量"，是指有一定长度、一定方向、一定位置的线段。这样汉字的笔画也就可以用数学上的矢量来描述。用矢量描述的汉字字模称为"矢量字模"，由矢量字模构成的字库称为"矢量字库"。矢量字库的优点在于，可以任意放大缩小，而且每个字模所占的信息量较小，整个字库容量可以大大压缩；缺点是不能直接用于输出。输出时还要把矢量字模转换为点阵字模。矢量字模的运用，更好地满足了对数字化字模高质量、高精度的要求。

汉字外码即汉字输入码，是为了把汉字输入计算机而编制的代码。汉字输入技术是向计算机传输汉字的交通工具，为了便于用户熟悉和掌握，常常要求输入码编码规则简单、易于记忆、操作方便、编码容量大、码长短和重码率低。

(2) 汉字输入

汉字输入方法可以分汉字键盘输入、汉字语音输入和汉字字形识别输

入三大类。

汉字键盘输入是利用键盘上数字键和字母键向计算机输入汉字的方法，汉字键盘输入法已有数百种，如周志农的自然码、陈代宇的大众码、朱邦复的仓颉码、王永民的五笔字型汉字输入法等。不同的汉字输入法的差异在于汉字输入码的不同编法上，可以分为字形编码类、字音编码类和音形结合编码类等。

字形编码类是以拆析的汉字字形为依据编成汉字的输入码，这种输入码叫"形码"。比如使用较为普遍的"五笔字型"的"王码"就属于这一类。这种输入法以汉字字形结构为基础，将笔画、部件、字形有机地结合起来。笔画共 5 种：横、竖、撇、点、折，部件是由笔画构成的用于组成汉字的构字单位，字形分 3 种：左右结构为一型，上下结构为二型，其他为三型。依照该方案的汉字编码规则，即可向计算机输入汉字。经过 6 次革新，这种输入法在输入速度、规范性等方面有了很大的改进，具有输入速度快、重码率低等诸多优点，成为使用范围最广的汉字输入方法之一。特别是新推出的"五笔数码"汉字输入法，利用数字键 1～6 或 1～9，可以实现单手输入汉字简体、繁体、香港汉字及词汇，易学易用，将在汉字计算机输入法的普及方面显示其优势。

字音编码类是以汉字的读音为依据的一种汉字编码方法，这种输入码叫"音码"。可以用汉语拼音(或再标调)编码，也可根据声韵形成双拼音标调进行编码。如智能 ABC、微软拼音输入法等都属于这一类。这类编码的特点就是编码规则与汉字音素有关，只要有一定的汉语拼音基础就很容易掌握，有较好的使用基础。但由于我国的方言较多，生僻字和同音字等情况复杂，因此这类编码在某些方面的使用会受到限制，如重码字较多，不会读的字无法输入等。

音形结合编码类是综合考虑汉字的形和音来编成汉字的输入码，根据编码时主要依据字音还是字形，又有"形音码"和"音形码"的区别。

在"字处理"的开始阶段是"字输入"，即一个字一个字地输入，这样的输入速度是比较慢的。经过不断改进，汉字输入已经进入到以字为基础、以词为主导、可以智能处理的"词组输入"或"句输入"阶段。这样一方面可以缩短输入码的长度，另一方面也可以缓解"重码"的困扰，从而大大提高了汉字输入的速度。不同的汉字输入方法各有优劣，"形码"速度很快，但因为要记住许多符号，所以适合于专业录入人员使用，不太适合记忆力逐步衰退的中

老年人。并且即使是年轻人，如果隔一段时间不用，也容易生疏。“音码”可以边想边打，但前提条件是要熟练掌握汉语拼音，在使用上也有它的局限性。因此，理想的汉字输入方法应该是“语音输入”或“汉字书写输入”。

汉字语音输入是无需借助键盘、直接由人口述输入汉字的方法，如IBM公司开发的Via Voice输入系统(非特定人连续中文语音识别系统)。该系统设置了输入者的口音计算机适应程序，计算机根据使用者对50～200多个句子的发音，建立起该用户的个人语音模型。另外该系统有一个包括数万条词的词表，从分词入手，看一段连续的语流所代表的汉字如何组合最合适，以此来提高输入的正确率，据介绍该系统的输入正确率可达92%。但也有实验表明该系统对同音词或近音词的识别正确率仍有待提高，如将“硕士”错成“设施、摄氏、特使”等。再如“能占(冷战)以来，/有一次，一个北大女生去清华的五会(舞会)跳舞，/并以(已)举行两次会议，/这女生惊呆了，以致一为(以为)他有什么机芯(居心)，/表情更苏木(肃穆)/”。所有括号内的词语都在该软件的“候选集”中出现，也就是说，系统选择的是最优解，但不一定是正确解。要解决这类问题，需要寻求用语言学知识/语言规则来发现和纠正错误。这样的分析在日语输入中也同样得到证明。其实，有关专家认为，语音技术的正确率目前至多达到80%，语音语料库语料的语言模型仅限于新闻体，日常口语的对话体还很缺乏，诗词更谈不上。虽然汉字语音输入是一种极为方便的方法，但仍有许多问题需要解决。

汉字字形识别输入是通过图形扫描仪器对汉字文本进行扫描，以使汉字输入计算机，实现扫描仪跟计算机对接的一种方法，包括印刷体汉字自动识别和手写体汉字自动识别两种。前者是国内汉字字形自动识别的主流，基本技术已经达到实用化阶段，但这种方法只能输入已经印刷好的成品，无法输入正在创作中的作品。

手写体汉字自动识别分为联机和脱机两种。联机手写体汉字自动识别是由使用者在一块跟计算机相连的写字板上写汉字，计算机同时接受笔的运行轨迹并进行识别输入。脱机手写体识别是识别预先书写在纸上的文字。手写体汉字识别的效率不仅与计算机识别技术本身有关，也跟书写者的书写规范有关，规范清晰的字形其识别的成功率总要高一些。

经过20世纪70～80年代我国科学家和工程技术人员的艰苦努力，中文信息处理中的第一道难关“字处理”已基本得到解决，无论是汉字输入技术还是汉字的计算机存储和输出技术都已进入成熟的实际应用阶段，特别

是在汉字键盘输入方面，输入速度甚至已经超过了英文字母的键盘输入，据测算，汉字键盘输入的效率可以达到英文输入的 1.3 倍到 1.9 倍。随着华光、方正等电子出版系统的开发成功和投入使用，汉字最终告别了铅与火的时代，这正像有人所说的，聪颖的华夏祖先创造了汉字和中华文化，智慧的炎黄子孙解决了汉字快速输入计算机的世纪性难题。

中文信息处理技术，是自然语言信息处理的一个分支。它需要以大量的语言知识、背景知识为依据，对中文信息的人脑处理过程进行模拟。当前的中文信息处理系统，还处于初级阶段，许多处理过程，尤其是机器翻译、人机对话等，还需对汉语本身加以研究，附加某些必要的选择条件和限制。例如，词汇限制，语句形式限制，语义和语用知识的限制，等等。

**二、国内中文信息处理的发展**

在世界上第一台电子计算机问世，被广泛运用于数值运算的同时，学者们也想到了利用这一工具自动地把一种语言翻译成另一种或几种语言，机器翻译成为 20 世纪 50～60 年代自然语言处理的中心课题。由于处理技术过于简单，未能取得预期的效果，上个世纪 60 年代中期，学者们转入了对自然语言的语法、语义和语用等基本问题的研究，尝试让计算机来理解人类的自然语言，研究领域不断扩大。

中文信息处理的研究在我国已有近 50 年的历史。早在 1956 年，我国学者就提出了研制中文电动打字机以及用汉字编码方法把汉字转换成信息代码进行传输等问题，引起了国内外学者的关注。1958 年，新华社、邮电部、中央机要局合作设计了鼓轮式中文电传机。1959 年国庆前夕，中国科学院计算技术研究所和语言研究所合作，在我国试制的第一台电子计算机上进行了俄汉机器翻译试验，这是中文信息和计算机的最早结合。1969 年 9 月，邮电科学研究院试制成功我国第一台电子式中文电报快速收报机，揭开了用计算机技术处理汉字信息的序幕。

为了促进中文信息处理研究工作的发展，1978 年在青岛成立了全国汉字编码研究会，这是我国中文信息处理方面的第一个学术团体。1981 年 6 月 27 日，在天津成立了中国中文信息研究会，后来改称中国中文信息学会，下设基础理论专业委员会、汉字信息处理专用设备委员会、汉字信息处理系统专业委员会、自然语言处理专业委员会、汉字编码专业委员会、计算语言学专业委员会、少数民族语言文字信息处理专业委员会等。在我国中文信

息处理蓬勃发展之时，欧美和日本也有不少学者研究中文信息处理，美国有中文计算机学会，日本有信息处理学会，都进行中文信息处理的研究。

20 世纪 80 年代以后国内学者不断将国外有关自然语言处理的理论和方法引进、介绍到国内，结合汉语的特点，探讨中文信息处理的有关理论问题。从语言知识抽取的角度看，众多的语法理论大致可以分为两类：一类侧重从语言事实中发现范畴，建立规则。像美国的描写语言学理论（结构主义语法）、法国特尼埃尔的依存语法（配价语法）、菲尔摩的格语法、以韩礼德为代表的系统功能语法、Langacker 倡导的认知语法等都属于这一类。另一类侧重如何将已发现的语言知识用一定的形式化方式加以描述。从乔姆斯基转换生成语法理论到后来与之相关的一系列语法理论，如扩充转移网络（ATN）、支配约束理论（GB）、功能合一语法（FUG）、词汇功能语法（LFG）、定子句语法（DCG）、中心词驱动的短语结构语法（HPSG）、广义短语结构语法（GPSG）、范畴语法（CG）等等，都属于这一类。当然，这些理论本身并不一定能给汉语语法研究带来更丰富的有关汉语结构规律的知识，但因其视角的新颖和描述手段倾向于严密的形式化表述，可以促进我们认识上的清晰化和对汉语语法研究的规范化。

80 年代中后期开始，学者们将国外的语法理论方法全面系统地应用于中文信息处理的相关理论探讨，陆续有学者将有关自然语言处理这一领域研究的基本理论方法加以总结形成系统。越来越多的语言学家汇入这一领域，一改传统以人为服务对象的研究方法，以中文信息处理为对象发掘、总结语言规则，从语言学角度分析中文信息处理的底层问题。

这一时期，学者们在理论探讨的同时，研制开发了一些中文信息处理的应用系统，基本解决了中文"字"处理的难题，出台了"信息处理用现代汉语分词规范"（中国国家标准 GB13715），开发了基本达到实用要求的中文自动分词以及自动词性标注系统，为全面深入地进行与自然语言处理相关的应用开发如机器翻译、人机对话等打下了基础。

为了给计算机提供更多的、可以利用的语言学知识，学者们致力于语言知识库的建设和汉语语法规则的发掘。鲁川、张普、林杏光等从动词与名词的语义关系入手，倡导基于格语法的汉语语义格关系研究，开发了供计算机使用的《动词大词典》，提出了一个由两大类、七小类共 22 个格组成的汉语格关系系统，根据与动词所表明的动作或状态相关的主体、客体间的语义关系，将动词分为他动词（如"吃、重视"）、自动词（如"走、毕业"）、外动词（如

"知道、懂")、内动词(如"病、死")、领属动词(如"有、具有")和系属动词(如"是、等于")等6个次类。在此基础上对1 000多个现代汉语常用动词按义项进行了动词语义格关系的描写,并对每个动词的各种格框架给出相应的例句来说明,对汉语动词跟名词性成分的语义搭配进行了概括描写。

朱德熙、陆俭明、俞士汶等以朱德熙提出的词组本位语法体系作为设置各项语法范畴的理论基础,研制开发了《现代汉语语法信息词典》①。选取一些具体的功能标准确定了汉语的词语分类系统,并对照一个词语的句法功能表现按义项把它归入某个词类;然后是以功能理念指导词语语法属性项目的设置,并根据一个词语的实际用法情况标记它的属性值。例如作为研究重点的动词在词典的总库和分库中共设立了100多项属性,来标记一个动词能否重叠、能否直接受名词修饰、能否作"有"的宾语、是带体词宾语还是带谓词宾语等等。根据每个词在这些属性项目上的取值,就大致可以确定一个词在实际的汉语话语中出现时的分布状况。

从抽取汉语语言知识的角度看,这两项工作在建立汉语语义知识和语法知识的基本范畴体系方面分别都做了相当深入的探索工作,但它们并不直接属于信息处理领域的某个具体的应用系统,而是出于通用的考虑,可以看作是在面向汉语的信息处理这个大目标下展开的构筑基础平台的工作。这些汉语知识库的建立,主要还是句法语义范畴的设置,缺乏直接服务于信息处理的汉语句法规则的总结。在这方面马希文、冯志伟、白硕、罗振声等进行过一些有益的探讨。要让计算机能正确判断词语之间的组合格局和内部关系,必须建立在对汉语的短语组合规则进行充分详细的描写的基础上,陆俭明、马真、孙宏林、詹卫东等面向中文信息处理展开了针对汉语某些具体问题的研究,力求在已知词项自身的范畴属性值的基础上,来发现多个词项在组合时的相互制约条件,或者是根据语符串的上下文环境来判断一个语符串的内部层次和关系。

随着中文信息处理的复兴,在国内还出现了基于不同理论背景的信息处理系统和方法,如董振东提出的"知网"和黄曾阳提出的以语义表达为基础来对汉语进行理解的概念层次网络理论(HNC)。当然,从全新的视角来探索中文信息处理的新途径,这些处理方法仍有很多工作要做。

此外,大规模语料库的建设及语料库方法的实际运用也是20世纪80

① 俞士汶等.《现代汉语语法信息词典详解》.清华大学出版社,1998

年代以后中文信息处理中的一个亮点，相继建成了一定数量的大规模语料库，开发了语料库自动检索、查询、标注等自动加工的工具，为语法规则的发现、统计信息的利用等提供支撑。

## 三、中文信息处理面临的挑战

### 1. 计算机处理语言信息的程序和基本方法

计算机对自然语言的处理，一般要经过如下几个阶段：

第一，从语言学的角度提出自然语言处理的问题和理论；

第二，把需要研究的语言学问题加以形式化，使之能以一定的数学形式或者接近于数学的形式，严格而规整地表示出来；

第三，把这种严格而规整的数学形式表示为算法，使之在计算上形式化；

第四，根据算法编写计算机程序，使之在计算机上加以实现。

把自然语言作为输入在计算机中进行处理时，在具体的实现策略上可以有不同的选择。如早期人机对话系统采用的简单模式匹配方法，后来发展起来至今仍在广泛使用的基于规则的处理方法，以及近年来日趋流行的基于大规模语料库的统计方法等。无论是哪种方法，都可以抽象为两部分来看，一是关于自然语言的知识，一是表述知识的机制。假定有关自然语言的知识是客观的，那么知识本身对规则方法和统计方法应该是共同的。这样，比较规则方法和统计方法的差异，就可以归结为表述知识的机制的不同。一般而言，规则方法最常见的是以一定的形式语法系统来表述自然语言中大小成分间的组合规则；统计方法则以各种统计数据来显示语言成分间的组合可能性。

计算机处理语言的情况与人学习语言有很多相似之处，计算机的计算是机械的，要靠人对它作出两个基本的规定：一是“做什么”，二是“怎么做”。计算机执行任何操作可以归结为一个公式：条件—动作，即在一定的条件下，执行一定的动作，在另一条件下，执行另一动作。人可以根据某种环境采取某种行动以及根据某一前提作出某种结论，这种现象与“条件—动作”的对应关系极为相似。要把人类的语言这种复杂的智能活动形式化，最关键的是要为某种语言建立一系列“条件—动作”偶对。要使人们更容易地学习或掌握一种语言，最根本的也是要告诉人们根据特定的条件来运用语言中的各种规则。要使自然语言的语法规则成为可供计算机执行的形式，必

须指出各种语法现象出现的条件。自然语言处理最根本、最关键的问题，是要指出各种语言形式出现和变换的条件。只有指出了条件，计算机才可能根据有关条件，执行相应的动作，从而使整个系统成为一个可以动态的执行的过程。要使自然语言的语法规则成为可供计算机执行的形式，就必须指出各种语法现象出现的条件。事实证明：凡是成功的处理方法都是由于比较充分地研究了语言形式出现的条件，凡是失败的，则或者是由于根本没有提出出现的条件，或者是虽然提出了语言形式出现的条件，但是条件给得不具体、不精确，或者是条件给错了。由于汉语中词类与句法成分、句法结构与语义结构之间的非一一对应性，在汉语自然语言处理中，研究各种“条件一动作”之间的关系显得尤为重要。

2. 中西方语言的差异对中文信息处理的影响

我们日常工作的信息绝大多数是以语言文字作为媒介来传播、交换和记载的。随着计算机的推广应用，由数据处理、信息处理发展到知识处理，对语言文字处理的要求无论从深度和广度上来说，都越来越高。这个问题在西方国家虽然也存在，但并不突出，特别是在字处理和词处理阶段。因为计算机从诞生之日开始，就是以处理西方语言为基础的。但是，汉语与西方语言差别很大，能够处理西方语言的计算机，面对汉语有时却显得无能为力。中文和西方语言文字上的差异给中文信息处理带来了诸多障碍。

1）从记录语言的文字体系上看，西方语言采用拼音文字，字符数量较少，如拉丁字母仅 26 个。记录汉语的汉字是语素文字，字符数量大，通用汉字有 7 000 多个，总数超过 5 万。文字类型的差异是中文信息处理中面临的第一个解决的难题，仅这一差异就引起了处理汉语的计算机与处理西方语言的计算机之间的一系列差异，存在着许多需要克服的困难，包括汉字的键盘输入、汉字的存贮、显示与输出以及汉字的自动识别等。

2）记录西方语言的文字实行分词连写，书面语的词与词之间有自然间隔，词界清楚。汉字记录汉语时不实行分词连写，字与字之间只有自然间隔，但不表示词界，因此词与非词的界定、中文的自动分词等问题是中文信息处理过程中面临的又一大难题。

3）从语言类型上看，西方语言有比较丰富的形态变化，外显的形态可以作为自然语言信息处理的客观依据。而汉语是典型的分析性语言，缺乏严格意义的形态变化，计算机对汉语的处理无法利用形态这一显性形式，只能通过隐性形式，在语法、语义上寻找出路。

3. 汉语研究传统与中文信息处理

自然语言的信息处理具有多领域、跨学科的性质，特别需要计算机科学和语言学的密切结合，要充分运用长期积累下来的语言学研究成果。在自然语言的信息处理研究中，如果没有语言学知识的支撑，再高明的计算机算法也难以发挥其应有的作用。语言学研究的成果直接推动了计算机技术的发展，语言学研究的进程也直接制约着计算机技术发展的进程。

计算机的参与，不仅更新了语言研究的手段，也向语言学提出了新的要求。传统的汉语研究主要是服务于人与人的交际，而现在必须考虑人与计算机的对话。人理解语言可以凭借相关的背景知识，可以凭语感判断语言现象，而计算机只能根据规则，对语言进行理性的分析，一步一步地推演和计算。因此，面向信息处理的语言研究不是罗列无穷的语言现象，而是要提供给计算机一系列可执行的条件，计算机需要的是形式化的语言规则，这些规则要具备可操作、可计算的特点。

我国的语言研究虽然具有悠久的历史，但长期以来的汉语研究多侧重于汉语教学，研究成果服务的对象是人，而不是计算机，所总结的有关汉语的规则有些无法直接在中文信息处理中运用，需要改造、深化、量化，甚至要从头开始。如对现代汉语“把”字句的认识，传统上认为“把”字句的作用是把及物动词的宾语提前，“关上门”可以说成“把门关上”，但将这一规则输入计算机，结果形成通篇的“把”字句。研究者发现，“把”字句还有其他一些条件限制，如动词为非光杆形式，“把”引进的对象是动词的受事且指称上是定指的等。再进一步概括：凡是受事主语句的主语前，都可以加上“把”形成“把”字句。规则更有效，更适合于在计算机上实现。

汉语的句法和语义两个平面相互交叉，又都与词汇义密切相关，因此汉语语法研究不仅是首要的而且也是一项复杂的工作。句法研究语言结构之间的关系，语义研究语言结构与现实世界之间的关系。句法与语义之间的相互制约表现为一种句法关系可以对应于多种语义关系，一种语义关系也可以通过多种句法关系表现。句法形式有限，而语义内容无限，以有限的形式表达无限的内容，必然产生一对多的结果，这样歧义的消解成为中文信息处理的一个重要方面。从计算机处理自然语言的角度看，自然语言中充满了歧义。如同音字和同音词问题，词处理中的歧义切分和兼类词问题，句处理中的短语边界的确定以及多义短语的消解问题等，对汉语研究提出了严峻的挑战。在自然语言处理中，歧义或多义是一个无法回避的问题，成为自

然语言处理的巨大障碍。如汉语的主谓结构中主语与动词之间的关系可以是施事—动作(朋友回国了)、受事—动作(书看完了)、结果—动作(坑挖好了)、处所—动作(北京去过了),动词所带的宾语可以是受事(踢球)、施事(来了一个人)、结果(写文章)、工具(乘飞机)、方式(跳街舞)、处所(去北京)等,甚至还有一些宾语,其语义角色难以认定,如"吃食堂"等。这种现象实质上是意义与形式之间矛盾的表现,同一形式与不同的意义相联系,必然产生歧义,这是自然语言不同于人工语言的特点之一。虽然此类现象对人而言未必带来理解上的障碍,但却是语言信息处理中必须面对的问题。这些问题的解决主要依赖于语言研究提出歧义消解的条件。以语言信息处理为目的的语言研究,不能仅满足于作句法上的分析,还要对这种结构所代表的实例的表义情况作出具体的分析,而且要做到可形式化,这样才能为计算机所应用,比如满足什么条件会出现什么结构等。

## 第二节 中文自动分词和自动词性标注

### 一、中文自动分词概述

#### 1. 自动分词和自动词性标注的含义

自动分词和自动词性标注是中文"词"处理阶段的两大主要任务,也是中文信息处理需要大力突破的关键技术之一。所谓"自动分词"就是让计算机确定中文文本中词与词之间的界限。记录西方语言的文字词与词之间有作为词界的空格,词与词之间的界限是清楚的。而记录汉语的汉字不实行分词连写,词与词之间没有空格,什么算是一个词,会引起极大的争论。可以这样设想,如果把某个英语文本中的所有空格符都去掉,再让计算机自动恢复作为英文词界的空格符,这就是词的识别或自动分词的过程。说得通俗一点,汉语的自动分词就是让计算机在中文文本中词与词之间自动地加上作为词界的空格符。

自动分词的实现可以说是让计算机知道了作为"词"这一级语法单位的知识,但由于汉语没有形态变化,汉语里的词哪个是名词、哪个是动词、哪个是形容词,无法依据显性的形态来判定。加之汉语里词类与句法成分之间存在一对多的关系,比如动词、形容词的主要功能是作谓语或谓语中心,但

它们也可以以相同的形式出现于主语、宾语、定语、补语、状语等句法位置上，这样如何确定汉语中某一个词的词性是“词”处理中的又一难题。所谓“自动词性标注”就是要让计算机来给切分出来的词自动地赋上相应的词类标记，为进行更高层次的中文信息处理提供语言知识做准备。值得注意的是，“分词”也好，“词性标注”也好，不是让人先把词分好，再把每个词的词性标注好，然后把真实文本输入计算机，而是让计算机自动完成。这样来看，所谓的“词处理”其实就是指让计算机面对真实的汉语文本自动地进行“分词”和“词性标注”的工作。

2. 中文自动分词的必要性和可能性

词是最小的、能独立运用的语言单位。计算机的所有语言知识都来自机器词典(给出词的各项信息)、句法规则(以词类的各种组合方式来描述词的聚合现象)以及有关词和句子的语义、语境、语用知识库。中文信息处理系统只要涉及句法、语义，就需要以词作为基本单位。分词以后在词的层面上进行处理，会大大提高处理的确定性。在更高层次的文本处理上，如句法分析、语句理解、自动文摘、自动校对、自动分类和机器翻译等领域，更是少不了有关词的详细信息。自动分词是任何中文信息处理系统都难以回避的第一道“工序”，其作用是怎么估计都不为过的。只有解决了这个难题，中文信息处理才称得上初步打上了“智能”的印记。如就中文文本的自动检索而言，如果不分词，当检索“华人”一词时，就会把“中华人民共和国”也检索出来，这样就会出现大量的检索垃圾，严重影响信息的采集。

由于长期以来学者们在有关汉语“词”这一级语法单位的大小问题、异同问题上存在着分歧，对于中文自动分词的可能性问题，尚不能下一个完全肯定的结论，但经过十几年的探索，中文自动分词已初具轮廓。毕竟“词”平面上的研究与句法平面和语义平面的研究相比较，难度要小得多，并且研究成果也更成熟和丰富，目前虽未能完全解决分词的问题，但已有一些汉语自动分词系统投入使用，并且可以达到比较高的正确率。

3. 中文自动分词所面临的困难

在过去的十几年里，汉语自动分词工作虽然取得了很大成绩，但与实用的需要相比，还有很大差距，还面临着诸多的困难。

(1) 分词规范

汉语的语素和词、合成词与短语之间没有明确的界限，要解决自动分词问题，首先要给计算机确立“词”的概念让它知道什么样的单位是“词”，什么

不是“词”,这方面的知识得由人来提供,这套知识实际上就是一个供计算机识别“词”的词表。词表分为基本词表(核心词表)和辅助词表。前者是为各行各业用户服务的通用词表,凡在该词表中的词,分词时就应该切分出来,但对于哪些词应当收进核心词表,目前尚无合理的可操作的理论和标准。后者作为前者的补充,是为特定用户服务的。分词结果是否正确需要有一个通用、权威的分词标准来衡量。分词标准的问题实际上是汉语词与语素、词与词组的界定问题,这是汉语语法的一个基本、长期的和有争议的问题。此外还有词的变形结构问题,如相对于“打牌”、“开心”、“看见”、“相信”而言的“打打牌”、“开开心”、“看没看见”、“相不相信”等的处理,缺少可操作而又合理的规范。再比如现代汉语中的词缀的问题,它们是构词要素,不能独立运用,如“记者、作者、拓荒者、成功者”等只能作为一个词看待,而“作出了巨大个人财产和精神牺牲者”、“克服许多困难而最终获得成功者”、“开发中国第一个操作系统软件者”等,无论如何是不能作为一个词看待的。

不同的应用需要,对词的切分规范有不同的要求,如以词为单位的键盘输入系统,为了提高输入速度,一些互现频率较高的相互邻接的几个字也常作为输入的单位,如“这是、每一、再不、不多、不在、这就是、也就”等。同样,中文自动校对系统、繁简自动转换系统、语音合成系统以及中文自动检索系统对词的切分也都有不同的要求。

(2) 切分歧义

汉语文本中含有许多可作不同切分的歧义字段,典型的有交集型歧义字段和组合型歧义字段。只有向分词系统提供进一步的语法、语义知识才有可能得到正确的处理结果。如“真正在”可以有“真/正在”或“真正/在”两种不同的切分可能性,“学生会”既可能是一个词,指一种学生组织,也可能是两个词“学生”和“会”。“才能”在“没有出众的才能就无法在竞争中站稳脚跟”是一个名词,而在“掌握新技术才能立于不败之地”是两个词,“才”是副词,“能”是动词。此类现象必须要借助于上下文、相关的句法语义信息以及频率统计等方法才有解决的可能。

(3) 未登录词的识别

按照一般的设想,似乎可以在计算机里建立一个庞大的词库,给每一个词都标注上一定的句法、语义信息,这样计算机在处理真实文本时,每碰到一个语言片段都先跟词库里所储存的词进行匹配检索,匹配得上的语言片段计算机才能识别,也就可以把那个词分出来。但计算机里的词库再大,也

不可能把实际语言里可能有的词都尽收无遗，再者汉语里词跟词之间的界限本来就不很清楚。因此现实交际中大量的人名、地名、机构名等，绝大部分都不会存储在计算机的词库里，计算机在自动分词时碰到这些词库里没有的专有名词就会束手无策。另外语言中的新词在不断地出现，这些新词在计算机词库里当然也是没有的，计算机在自动分词碰到这些新词时也会感到无奈。所谓未登录词指未收录在词表中、但必须切分出来的词，主要包括各类专名如人名、地名、企业名、商标名等，以及某些术语、缩略词、新词、外族和外国名的汉译名等等。如“于大海发明爱尔肤护肤液”需要切分成“于大海/发明/爱尔肤/护肤液”，并需要计算机识别出“于大海”是人名，“爱尔肤”是商标名，“护肤液”是术语名词。

4. **中文自动分词的方法**

即使有一个比较完善的分词词表，计算机在对一串连续的汉字字符进行切词时，也常常会出现使计算机举棋不定的现象，造成切词的失败，或者得出错误的切分结果。因此自动切词方法对分词的正确率有着重要影响，目前国内采用的自动切词方法大体上可以分为基于字符串匹配的分词方法、基于理解的分词方法和基于统计的分词方法几类：

(1) 基于字符串匹配的分词方法

这种方法是按照一定的策略将待分析的汉字串与一个“充分大的”机器词典中的词条进行匹配，若在词典中找到某个字符串，则匹配成功，识别出一个词。按照不同长度优先匹配的情况，可以分为最大(最长)匹配和最小(最短)匹配；按照扫描方向的不同，可以分为正向匹配和逆向匹配；按照是否与词性标注过程相结合，又可以分为单纯分词方法和分词与标注相结合的一体化方法。还可以把几种方法结合使用，如双向扫描法，这种方法分别用最大匹配法进行正向和逆向的扫描和初步的切分，并将用两种方法初步切分的结果进行比较，如果两种结果一致，则判定切分正确，如果两种不一致，则判定为疑点。这时，或者结合上下文有关信息，或者进行人工干预，选取一种正确切分。

(2) 基于理解的分词方法

通常的分析系统，都力图在分词阶段消除所有歧义切分现象，而有些系统则在后续过程中来处理歧义切分问题，其分词过程只是整个语言理解过程的一小部分。其基本思想就是在分词的同时进行句法、语义分析，利用句法信息和语义信息来处理歧义现象。它通常包括三个部分：分词子系统、句

法语义子系统和总控部分。在总控部分的协调下，分词子系统可以获得有关词、句子等的句法和语义信息来对分词歧义进行判断，即它模拟了人对句子的理解过程。这种分词方法需要使用大量的语言知识和信息，由于汉语语言知识的笼统、复杂性，难以将各种语言信息组织成机器可直接读取的形式，因此目前基于理解的分词系统还处在试验阶段。

(3) 基于统计的分词方法

从形式上看，词是稳定的字的组合，在上下文中，相邻的字同时出现的次数越多，就越有可能构成一个词，因此字与字相邻共现的频率或概率能够较好的反映成词的可信度。可以对语料中相邻共现的各个字的组合的频度进行统计，计算它们的互现信息。互现信息体现了汉字之间结合关系的紧密程度，当紧密程度高于某一个阈值时，便可认为此字组可能构成了一个词。这种方法只需对语料中的字组频度进行统计，不需要切分词典，因而又叫做无词典分词法或统计取词方法。但这种方法也有一定的局限性，会经常抽出一些共现频度高、但并不是词的常用字组，如“这一、之一、有的、我的、许多的”等，并且对常用词的识别精度较差。实际应用的统计分词系统都要使用一部基本的分词词典(常用词词典)进行字符串匹配分词，同时使用统计方法识别一些新的词，也就是将频率统计和字符串匹配两种方法结合起来，既发挥匹配分词切分速度快、效率高的特点，又利用了无词典分词结合上下文识别生词、自动消除歧义的优点。

## 二、歧义字段与自动分词

### 1. 交集型歧义字段

交集型歧义字段简称交集字段。假设 A，B 和 C 分别表示一字或多字组成的字串，如果在 ABC 字串中，A，AB，BC，C 都是词表中的词，则 ABC 称为交集型歧义字段。如“太平淡”中“太平”和“平淡”分别成词，“太平淡”是交集型歧义切分字段，“平”为交段。再如“应用于”、“可以为”等都属于这一类。交集字段中相互交叉的词叫做交集因子。如“应用于”中“应用”和“用于”都是它的交集因子。一个交集字段中交集因子的数量称为这个交集字段的链长。如“应用于”的链长为 2。

针对中文自动分词中的交集型歧义字段，山西大学的郑家恒、刘开瑛通过一个 180 万字次的新闻语料库中交集型歧义字段的普查，共取得了 9 500 条交集型字段，并据此建立了一个交集字段库，发现在所有的实例中以 2 字

为一切分单位的字段达 8 378 条(占 88.2%),而链长为 2 和 3 的交集型歧义字段占总数的 96.1%,它们的出现次数则占总次数的 97.7%。因此如果能解决好链长为 2 和 3 的交集字段的切分问题,就可以大大提高整个歧义字段切分的正确率。如对链长为 2 的交集字段 ABC 而言,切分结果有如下 4 种情况:

① 切分结果为 A/BC,如"出/自己";

② 切分结果为 AB/C,如"出现/在";

③ 切分结果为 ABC,如"传染病";

④ 切分结果不定。

统计结果显示,切分为前两种的交集字段占总数的 89.4%,出现次数占总次数的 80.6%,其中 AB/C 型在正向最大匹配法的切分过程中可以得到正确的切分。因此对链长为 2 的交集字段,切分难点集中在 A/BC 型交集字段的辨识上。

同样,对链长为 3 的交集字段 ABCD 中切分结果有:A/BC/D, A/B/CD, AB/CD, ABC/D, AB/C/D, ABCD 以及切分结果不定这 7 种情况。统计表明,AB/CD 型交集字段无论在交集字段的个例数还是在出现次数上都占总数的约 98%,如"已经过去"。因此对链长为 3 的交集字段,一般来说切分结果应为 AB/CD。

清华大学的孙茂松等利用清华大学的词表 TH-WL,内含 112 967 个词条,在一个规模约 1 亿字次的新闻语料库 Rcorpus 上,对交集型歧义字段进行了研究,共搜集到歧义字段 233 888 条,通过分析,他们将交集型歧义字段分为 3 种:一种是伪歧义,即虽然有歧义的可能性,但在真实文本中通常只有一种切分结果;一种是真歧义 1,经常存在两种以上的切分结果;一种是真歧义 2,本质上属于真歧义,但通常也只有一种切分结果。也就是说,其他形式的切分结果出现的机会很少,基本上可以当成伪歧义处理。他们对其中 4 619 例高频歧义字段进行了统计分析,发现伪歧义占 92.6%,真歧义 1 占 1.9%,真歧义 2 占 5.5%。

交集型歧义字段是由词与词之间的交叉组合产生的。如:

用树形图形式加以描述(图/形式,图形/式);

研究生命的本质(研究生/命,研究/生命);

白天鹅游过来了(白/天鹅,白天/鹅);

让位移等于 50 厘米(让/位移,让位/移);

独立自主和平等互利的原则(和/平等,和平/等);
这本小说的情节太平淡了(太/平淡,太平/淡);
对这种现象的确切描述(的/确切,的确/切);
社会需求和生产水平有矛盾(需/求和,需求/和);
他们看中和日本人做生意的机会(看/中和,看中/和)。

在大规模语料库的基础上对交集型歧义字段的普查,对于有针对性地设计自动排歧方法十分必要,可以大大提高此类歧义字段的切分精度。

2. 组合型歧义字段

如果字串 AB 中,A,B 和 AB 都是词表中的词,则 AB 称为组合型歧义字段。如"阵风"中"阵"、"风"和"阵风"都是词表中的词,"烤白薯"中"烤"、"白薯"和"烤白薯"都是词表中的词,则它们都是组合型歧义字段。

组合型歧义字段比较复杂,这种歧义字段是由词与词之间的串联组合产生的。如:

这条马路可以并排行驶四辆大卡车。(副词)
教务科指定了专任讲师并排好了课程时间表。(连词+动词)
我一看他的穿着打扮就知道他不是等闲之辈。(名词)
她今天是穿着一身礼服出去的。(动词+助词)
你们可以对比一下这些布料的选色和构图。(动词)
他只对比他强的选手有兴趣。(介词+介词)
您过奖了,我做的只是我应该做的事情。(动词)
他参加过奥林匹克数学竞赛,还得过奖呢。(助词+名词)
你们后天再来吧。(名词)
到家后天就黑了。(方位词+名词)
每个人手上都拿着火把,照得黑夜比白天还亮。(名词)
这场火把他全部的希望都化为了灰烬。(名词+介词)
你还记得这件事情发生的时间吗?(动词)
记分的时候我并不是记得了多少分,而是记该扣多少分。(动词+动词)
他马上就来。(副词)
他从马上摔下来了。(名词+方位词)
这部电影的女主角还没有合适的人选。(名词)

三个人选一个，你会选谁？（名词＋动词）

他最近发了一笔小财，很是得意。（名词）

离我们家最近的那家餐馆失火了。（副词＋形容词）

可见，在组合型歧义字段中，歧义字段就是一个歧义词，而非歧义词被包含在歧义词当中。为了正确地切分组合型歧义字段，可以利用有关的句法信息、语义信息和语用信息。

句法信息：有些组合型歧义字段与其前后的字串之间具有密切的搭配关系，这时就可以利用有关的句法信息得到正确的切分。如"阵风"由量词"阵"和名词"风"组合产生，按非歧义切分时的词间搭配关系，量词之前应该有数词，因此，可以先在词法知识库中对歧义词"阵风"加上歧义标志并建立如下规则："如果歧义字段的前字串是数词，则歧义字段的首段单切，否则该歧义字段成词。"

语义信息：词与词之间的组合存在语义的相容性选择关系，即一个词对与之组合的其他词具有一定的语义要求，在自动切词过程中，可以有效地利用此类信息。如：

学生/会/兴奋/得/手舞足蹈。

学生会/兴奋/得/手舞足蹈。

这两种切分结果在词类与句法结构上十分相似，因此仅仅利用词法和句法知识，难以对这两种切分结果作出正确的判别，也就难以作出正确的切分。从语义上看，形容词"兴奋"的义项中要求其主体应该具有"人"这个义素，名词"学生会"的义项中不具有该义素，而名词"学生"的语义中有，利用这项语义知识，可以建立如下规则：如果歧义切分字段的后续动词/形容词中包含有动作发出者"人"这个义素，则歧义字段的尾字单切，否则该歧义字段成词。这样就可以得到正确的切分：学生/会/兴奋/得/手舞足蹈。

语用信息：如"乒乓球拍卖完了"仅根据词法、句法和语义信息都不足以判断卖完的东西究竟是"乒乓球"还是"乒乓球拍"，这时就得根据语言交际的具体环境的语用方面的知识，才能决定究竟什么才是正确的切分。

3. 歧义字段的消歧

十几年来，研究人员几乎调动了人工智能领域所有先进的计算手段来对付切分歧义，在一定范围内取得了各自的效果，但从总体上看，还都嫌粗糙。通过不断的实践，人们越来越深刻地认识到，如果没有足够的语言学知

识作为支撑，再先进的计算手段也只能是中看不中用。切分歧义消解经历了一个由浅及深、由简单到复杂的语言知识利用的演变过程。

早期的系统，主要利用词频以及语素、切分歧义表层结构等简单信息来消除歧义，有人揭示了音节信息在自动分词中的作用，有人断言，95.0%左右的切分歧义可以借助句法以下的知识解决，只有5.0%必须诉诸语义和语用知识。但规则往往是由人凭主观编制而成，会受到系统性、有效性、一致性、可维护性等问题的困扰。为克服人工句法规则集的弊端，也有研究人员尝试通过句法统计来消除歧义，提高分词精度。

当然，分析的层次越深，计算机系统对知识库的质量和规模等的依赖性就越强，所需要的时间、空间代价也就越大。孙茂松、左正平等发现，从一个1亿字真实汉语语料库中抽取出的前4 619个高频交集型歧义切分覆盖了该语料库中全部交集型歧义切分的59.20%，对另一个完全独立的语料库的覆盖率为50.85%，这说明高频交集型切分的分布相对不同的领域是比较稳定的，其中4 279个属伪歧义（如"和软件"、"充分发挥"、"情不自禁地"），覆盖率高达53.35%。由于伪歧义的消解与上下文无关，于是他们提出了一个简单却很有效的策略：对伪歧义型高频交集型歧义切分，可以把它们的正确（惟一）切分形式预先记录在一张表中，其歧义消解通过直接查表即可实现。

## 三、自动词性标注

### 1. 汉语的词类问题

所谓自动词性标注就是利用计算机来自动地给文本中的词标注词类信息。汉语书面文本在自动切分之后，词与词之间出现了代表词界的空格，这样就有可能像处理英语、法语、德语等语言那样，进一步分析每个词的词类，并给每一个词自动地标注上词类信息。在英语、汉语等自然语言中，都存在着大量的词的兼类现象，这给文本的自动词性标注带来了很大的困难。因此如何排除词类歧义，是文本自动词性标注的关键问题。

词类指的是词的语法分类，自汉语语法学建立以来，词类一直是困扰汉语语法学界的一个"剪不断，理还乱"的老大难问题，诸如汉语的词类划分标准问题，词类与句法成分之间的关系问题，都曾引起过极大的争议。经过100多年来的研究，人们认识到，词类的划分不是人强加给语言的，而是语言自身存在的特征，面向人的语法研究需要给词进行语法分类，面向计算机

的语法研究同样不能回避这一问题。

由于汉语缺少形态变化，词的语法身份是隐性的，中文信息处理作为一项庞大的系统工程，需要在计算机中建立若干种有关语言的知识信息，自动分词使计算机初步具备了有关汉语"词"的知识，虽然仍未达到完全成熟的程度，但已取得了令人鼓舞的成果。语言运用不是词的随便堆砌，要受到来自语义、语法和语用等方面规则的制约。自动分词的初战告捷，使我们有可能让计算机学习更多的有关汉语的知识，从而为深层次的中文信息处理打下基础。在用于自然语言处理的多种语法理论中，词类往往是作为元代码存在的，如转换生成语法中包含 N，V，A 等词类标记，用上下文无关方法分析句子所得到的句法树节点，也都是词类代码。因此词性标注是中文信息处理中的一项关键技术。

虽然在汉语词类划分的标准上存在过分歧，但事实证明，汉语的词完全可以分类。《现代汉语语法信息词典详解》在朱德熙词组本位语法理论的基础上，成功地对 5 万多词进行了词类划分，并且已被数十个单位的语言信息处理系统所采用。

词性标注实际上是给计算机增加了有关"词"的语法信息，信息处理用的现代汉语词类体系必须充分吸收有关现代汉语词类研究的语言学成果，结合信息处理的特点，设计合理的词类体系。《现代汉语语法信息词典详解》的词类系统包括 18 个基本类：名词、时间词、处所词、方位词、数词、量词、区别词、代词、动词、状态词、形容词、副词、介词、连词、助词、语气词、拟声词和叹词。此外，从计算机处理实际文本的需要出发，从提高处理效率的角度考虑，增加了一些不能视为"词"的单位的标记，包括比词小的单位和比词大的单位，如前接成分（阿、老、非、超、小）、后接成分（于、性、员、者、们）、语素（民、衣、物）、非语素字（蟋、枇、鸯、葡）、成语、习用语以及简称略语。上海师范大学在范开泰主持的国家汉办课题"当代汉语语料库"的加工中，设置一级词类标记 14 个，其他 9 个，共 23 个，二级词类标记 69 个，其他 7 个，共 76 个，优点是便于根据不同需要提取相应的语料。

*2. 汉语词归类的困难*

(1) 词类与句法成分的关系

撇开词类划分标准上的分歧，就以句法功能为标准对汉语的词进行语法分类而言，汉语的词类划分仍要处理好词类与句法成分之间的关系问题。如：

| | |
|---|---|
| 他来了 | 他在调查情况 |
| 来可以了解情况 | 调查很及时 |
| 他愿意来 | 这件事需要调查 |
| 来的时间不好 | 调查的结果很难说 |

同一个词可以出现于不同句法位置上,如何归类?也许有人认为,根据不同的句法位置标注不同的词性有助于消解歧义,减少自动分析的困难。其实不然。按照当前的处理技术,计算机进行自动分析的前提是要知道语句中词的词性,然后才能确定词在句法树中的位置及功能。如果词典中存在大量的多类词,在分析句子时就会遭遇非常严重的自动词性标注多选一的问题。事实上出现于不同句法位置上的同一个词仍具有大致相同的功能,如上例中的"来"在主语、宾语、定语位置上时仍体现出动词的特性,前面可以有状语、主语,也可以带宾语。如"经常来/他来可以了解情况"、"他愿意经常来"、"来中国的时机很好"等。汉语的词类具有多功能性,词类与句法成分之间不存在一一对应的关系。

(2) 兼类问题

兼类现象的界定和处理是汉语语法研究中的难点,当然也是自动词性标注的棘手问题。所谓兼类通常指同一个词在不同的语言环境中具有几类词的语法功能,"同一个词"的含义是同字、同音、意义相关。如"共同完成/共同愿望、自动控制/自动步枪、定期检查/定期存折"中的"共同、自动、定期"形和音完全相同,意义也没有太大的差别,但一个充当状语,一个充当定语,都不能充当其他句法成分,应视为兼属区别词和副词。

兼类问题是汉语词性自动标注中的一个难点,有点类似于自动分词中的歧义字段的状况。兼类是客观存在,但无论是对人还是对计算机,兼类不能太多,否则会落入"词无定类"的陷阱。

(3) 同形词

自动词性标注除了要处理好兼类词问题以外,同形词也是需要慎重对待的问题。同形词指的是词形相同而意义没有联系的一组词,在面向人的语法研究中,它们与兼类之间的界限比较容易确定。但中文自动词性标注是以书面汉语文本为对象的,计算机除了具备"词"的知识以外,没有过多的语言知识,如词的读音信息、词的意义信息等。同形而功能不同的词在自动词性标注过程中面临着与兼类相似的困难。如"两封信"与"别信他"中的"信"同形同音,但前者为名词,后者为动词;"在广大学生中"与"射中了"中

的"中"读音不同，词性也不同，前者是方位词，后者是动词。由于计算机处理文本通常不考虑语音，同形词和兼类词的识别处理没有本质的区别，这导致了词类标注中遇到的词类歧义现象比语言学意义上的兼类现象范围更广。包括以下几种情况：

字同音不同的同形异音词，如"好"，第 3 声是形容词，第 4 声是动词。

字同音同义不同，如"编辑"，指人是名词，指动作行为，是动词。

字同音同的兼类词，如"丰富"，作动词时，可带宾语，作形容词时，前可加程度副词。

多类混合体：如"了"在现代汉语中可以有动词、助词、语气词等多种性质。

可见兼类与同形这两种不同性质的现象在中文信息处理的自动词性标注策略上有许多共同性，因而有人将这些词形相同而功能不同的词统称为"多类词"。从广义上讲，多类词是存在于自动词性标注过程中的歧义问题。如：

着：助词、名词、动词、量词　　过：动词、副词、助词
就：副词、动词、介词　　下：量词、动词、方位词
两：量词、数词　　好：动词、形容词、副词
里：量词、方位词　　一：数词、副词、连词

通过对真实文本中多类现象的调查统计，多类词的数量不多，但使用频度较高；越是常用的词，多类现象越严重；多类词现象复杂，常见的包括：动词-量词、数词-副词-连词、介词-动词、动词-名词、量词-名词、动词-方位词-名词、副词-助词-动词、介词-量词、数词-动词、动词-副词等多类现象。

3. **多类词的排歧**

从大的方面看，自动词性标注中多类词的排歧可以分为基于规则和基于统计两种方法。

1）基于规则的处理方法是利用上下文框架规则描述在特定的语言环境下，一个多类词的词类标记，语境信息包括词语信息、词类信息、词语特征信息等。针对使用频率高的多类词可以设置特定的规则，检查它所出现的上下文环境，惟一地确定该词的词类标记。如针对多类词"了"可以设置如下规则：

"了"＋句尾标记：语气词，如"下雨了"。
动词＋"得|不"＋"了"：动词，如"吃得了、吃不了"。
动词|形容词＋"了"：助词，如"去了上海/脸红了"。

调查语料中经常出现的词类组合,分析它们在不同语境下选取某个词类的可能性,通过不同词类的语法功能特征,特别是某类词区别于他类词的特有的语法信息,构造相应的上下文规则来选择正确的词类标记。如"方便"类可以有动词和形容词两种词类标记,根据以下词类组合可以排除其多义性:

{"很/太/最/非常"……}+动-形:形容词,如"很方便/丰富/繁荣"。

量词+动-形+名:形容词,如"(一)种方便食品"。

动+动-形+名|助|语气:形容词,如"购买方便食品"。

其他情况:动词,"大大方便了群众"。

设置一些上下文词类标记匹配模式,描述在一定的词类环境下可能出现的词类标记集合,利用集合运算确定标记。如动词和名词之间可以出现"介词(p),助词(u),形容词(a)",对于"有/v 怪/ 招儿/n"而言,因为"怪"可以有"形容词(a),副词(d),动词(v)"3 种可能性,{p, u, a}∧{a, d, v}=a,可以确定例中"怪"的标记为"a",即形容词。

基于规则的方法要考虑两个问题:一是规则对语言现象的覆盖率,二是规则处理的正确率。一般而言,对于一条规则这两种性能往往呈反比关系,即如果一条规则覆盖的语言现象越多,则它处理的正确率越低;反之,规则覆盖的语言现象越少,则它处理的正确率越高。针对个别词的规则可以对有关信息描写得非常深入,因而正确率较高,但能处理的语言现象有限。

2) 基于统计的多类词排歧方法得先选取一定量的语料库作为训练集,对该训练集中的语料逐词进行人工标注,然后利用计算机对该集中的任意两个相邻标记的同现概率进行统计,形成一个相邻标记的同现概率矩阵,依据同现概率的高低确定多类词的词类标记。

这方面山西大学曾选取 10 万字的语料作为训练集,共采用 174 个标记进行人工标注。首先建立非兼类词词典,对于大量非兼类词通过该词典直接标注;其次建立兼类词词典,对每个兼类词除标记出各种可能的词类标记外,还标明各自在语料中的出现频率;第三,建立相邻标记的同现概率,形成"174 · 174"的概率矩阵;最后确定标记跨段,即以非兼类词为标界,如果其中包含兼类词,则作为一个标记跨段,根据所有可能途径的频率计算,依照相邻标记的概率,求出每条路径的总概率,最大的即为最佳路径,从而确定兼类词的词类标记。

基于规则的多类词性标注不必事先建立训练集，只要语言学家能提供足够的可以形式化的语言规则，就可以达到理想的排歧效果。但目前的语言规则还很难满足自动词性标注的需要，难以处理复杂的、不规则的信息。基于统计的方法所有的知识来自事先选定的训练集，计算机从中自动获取相关知识，具有较强的一致性和较高的覆盖率，可以将一些不确定的知识量化。但是这种方法获取的知识很难与现有的语言学成果结合，也很难用现有的成果来改进统计所获取的知识。目前用得比较多的是将两种方法结合起来，一方面通过语言学知识的引入，可以消除由统计方法得到的一些不合乎语言学规则的信息，又可以使由统计方法建立的语言学模型能很方便地与现有的语言学方法相结合。另一方面，由于统计方法的语言知识来自大规模的真实文本，覆盖面较广，能克服语言规则的片面性和主观性，可以集中精力研究最常见的、在统计意义上最重要的语言现象，从而从根本上提高汉语自动词性标记的正确率。

20世纪80年代以来，计算机学界和汉语言学界正在联手攻克中文"词"处理中的自动分词和自动词性标注这一难关。一方面通过进一步改进和完善分词规范，建立分词标准。另一方面组织科研力量，专门研究人名、地名的识别问题，专门研究计算机自动记忆新词、自动定称新词等问题。20多年来在中文"词"处理方面都已经取得了可喜的成果，也研制出了一些相应的软件。自动词性标注系统已投入实际使用，计算机自动分词的正确率也已达到90%～95%。当然，中文信息处理界也不会小看那剩下的5%～10%的分词问题，现在正在继续努力解决分词中的种种难题。自动分词和自动词性标注从某种意义上说是个关键性的技术关卡。词处理解决好了，既可以进一步反过来推动字处理技术的发展，而且也有助于"句处理"技术难关的攻克。

## 第三节　中文句处理

### 一、中文句处理概述

#### 1. 中文句处理的含义

中文信息处理在成功地越过汉字的输入、存储和输出的第一道难关，自

动分词和自动词性标注基本能满足实际运用需要以后，便进入了更高层次的阶段，即句处理阶段。所谓中文信息的“句处理”指的是，怎样使计算机处理、理解自然语言中的各个句子的意思，又怎样使计算机生成符合自然语言规则的各种句子。

说句处理是中文信息处理的更高层次的阶段，是因为这一阶段面临着更多、更复杂的问题，句子是由词、短语或小句构成的、具有交际功能的表述性单位，真正理解一个句子包括理解句子的结构、句子的意义以及句子的内容等方面，而要达到这一目标，涉及到以词语串形式出现的句子内部构成成分(短语)的结构定界、结构关系和外部功能的识别，涉及到相同功能类排列模式多义短语的排歧，涉及到不同层面的语义问题，如实词的自身义、实词之间的关系义、实词之间的组合义、特定的范畴义等，还涉及话题、焦点等语用因素对句子意义的影响。句处理已成为当前中文信息处理中最富挑战性、迫切需要解决的问题。

**2. 中文句处理的策略**

为实现这一目标，现在“句处理”有多种策略和途径：有基于句法规则的，有基于概念网络的，有基于语料库统计的，有基于语义计算的，等等，形成了一个竞相发展的局面。这些策略和途径表面上看各有侧重，其实都包含着两个方面的内容：一是关于汉语本身的知识，一是表述汉语知识的机制，两者都离不开汉语研究成果的支撑，最终都需要依赖可靠的汉语知识来指导计算机正确地处理自然语言。句处理所需要的汉语知识是涉及语音、语义、词汇、语法、语用等诸方面的综合性知识，目前有关汉语的知识还远远不能满足中文句处理的需要。许多在人看来还是比较容易解决的问题，但计算机可能解决不了。例如：

北京的公路建设得很快。

北京的公路建设很有成绩。

两句中字面上有相同的部分“北京的公路建设”之间的差别在人看来不难区分，前一句说的是“北京的公路/建设得很快。”后一句说的是“北京的公路建设/很有成绩。”但如果没有足够的知识，计算机可能将这一形同而实异的现象作相同的分析从而导致理解的失败。再比如：

企图逃跑、总结归纳、表演结束、推算出来、开门出去、派他出国、生产管理、讽刺说、介绍写(小说的经验)

上几例都可以抽象化为“动＋动”的词类排列模式，但符合这种模式的实例中，有些可以组合成合法的句法结构体，而这些合法的句法结构体内部成分即两个动词之间的结构关系不同，可以有动宾、联合、主谓、述补、连动、兼语、偏正等多种结构关系。也有些不能构成合格的句法组合，如“介绍写”。这类相同功能类成分的排列构成合法的句法组合的条件是什么，满足什么条件，该合法组合是何种结构类型和功能类型，目前语法学界对此还不能给出十分清楚的答案，更不用说让刻板地按照给定条件执行相应动作的计算机去处理了。计算机对汉语的句法分析尚且如此困难，处理汉语的语义问题、语用问题面临的困难会更多。

中文字处理特别是词处理的初步成功表明，作为一项系统工程，中文信息处理要取得比较理想的成果，必须坚持把基于统计的方法与基于规则的方法、句法知识与语义知识有机地结合起来。既需要建立基于语料库的统计分析，也需要有按汉语本体研究已获得的规则支持；既要注重句法分析，也要注重语义处理；充分借鉴和吸收汉语研究理论与国外有关理论、现代语言学理论与中国传统语言学理论的精华，把基础研究理论与应用系统的研制开发结合起来。这样才能使我国在不太长的时间里，在中文信息处理乃至自然语言理解和处理方面，从工程到理论，能达到世界先进水平，继而能居于世界领先地位。

虽然目前的中文句处理有不同的策略，但从大方面看，不外乎基于规则和基于统计两种。

基于规则的处理方法其特征是充分利用语言学的已有研究成果，整理、确立一定的范畴体系，在此基础上建立计算机进行“句处理”所需要的词库，在词库中对每一个具体的词或短语进行尽可能详细的属性赋值。然后，整理、确立能正确地描述范畴之间关系的规则，在计算机内建立相关的规则库。词库和规则库是计算机进行运算、分析的依据。研究者根据计算机处理的结果，做相应的调整工作，改进词库和规则库的内容，以提高处理的正确率。有关自然语言的知识可以分为两大类，一类是基于范畴的“属性：值”型知识，一类是基于规则的“条件→动作”型知识。“范畴用来刻画语言对象的一个或一组特征。规则用来表述范畴间的关系”①。范畴知识可以详细

① 詹卫东．面向中文信息处理的现代汉语短语结构规则研究．清华大学出版社、广西科学技术出版社，2000．3

地记录在机器词库中,规则知识则集中体现在规则库中,规则的作用在于建立起不同范畴之间的联系。要使基于规则的处理方法能取得比较理想的效果,首先要求语言学家能提供足够详细、可信、可形式化的语言知识。但由于以往的语言研究主要是面向人的,语言学家所提供的相关语言知识还只是粗线条的,难以满足计算机处理语言信息的要求,因而在实际工作中遇到了重重困难。

基于统计的处理方法是随着大规模语料库的陆续建成而出现的一种自然语言处理方法,其主要特征是基于计算机对大规模语料库真实文本的统计分析,由计算机来抽取语言知识。

语料库是存放语言材料的仓库,是以计算机为载体承载语言知识的基础资源。根据加工深度,语料库可以分为不同的等级,如 0 级语料库或生语料库,1 级语料库(完成分词和词性标注),2 级语料库(完成短语结构关系、结构功能标注)等。由于语料库是最真实、最可靠的语言知识源泉,而且都已经过不同程度的加工,真实文本中已负载了一定语言信息,为语言知识的抽取提供了极大的便利。比如可以在完成自动分词和词性标注的语料库基础上,进行双词语环境的语言知识抽取,这对于句处理中的短语结构定界、词语之间的搭配限制具有重要意义。基于统计或大规模语料库的句处理方法主要利用计算机存储的语料和计算机的高速运算速度,从语言使用的现实状况中由计算机自动归纳、总结语言规则,再利用得到的统计数据或结果对语料进行分析,再根据分析得到的反馈结果调整已有的规则,逐步提高语言处理的精度。

基于规则的和基于统计的句处理策略各有优势,也都存在着不足。基于规则的方法依赖语言学家对语言的研究成果,由人来对语言知识进行抽象;基于统计的方法依赖计算机对大规模语料库的统计分析,由计算机来抽象语言知识。目前越来越多的学者提倡将两种方法结合使用,使这两种方法的长处得到充分的发挥,取长补短,使两种方法甚至其他的方法能在中文句处理的不同方面发挥应有的作用,以期能在较短时间内取得较好的效果。

从目前的认识和中文句处理的进展情况看,实现中文句处理的策略是:

1) 规则尽可能少,少到最低限度;词库中每一个语言成分的信息尽可能详细充分。这与当代语言学界所倡导的"大词库,小语法"的思路是一致的。这样做的好处是,不至于因为某条规则的修改而全面改动中文信息处理的规则系统。

2）基于统计的方法和基于规则的方法相结合，既需要基于语料库的统计数据，也需要通过汉语本体研究获得的规则。

3）尽快为计算机建设、提供两种类型的数据库，一是具备一定规模、经过深加工的现代汉语平衡语料库，二是有关现代汉语的各种知识的数据库。

4）努力探求好的分析方法，既要注重句法方面，也要注重语义、语用、认知方面。

3. *中文句处理需要着力解决的问题*

中文句处理涉及多方面的语言知识，就目前来说下面几项工作与句处理的开展关系尤为密切：

1）汉语动词的题元结构系统。包括建立面向汉语动词的题元结构理论框架、不同动词的题元结构的形式化模型，不同的词语参与组合后可能发生的变化研究，以及对各个具体动词（按义项出条）进行动名句法形式选择和动名题元角色标注。

2）汉语词的语义分类层级系统。包括建立适用于汉语的语义特征分析的理论框架以及建立能与动词题元结构系统配套的现代汉语词汇语义层级系统。

3）词语之间的组合类型与组合规则。包括同类的词与词之间彼此组合可能形成的种种关系类型及其具体规则的研究、非同类的词与词之间彼此组合可能形成的种种关系类型及其具体规则的研究，以及同类的或不同类的短语与短语之间彼此组合可能形成的种种关系类型及其具体规则的研究。

4）句法、语义的多功能性研究。从功能范畴的角度设想并研究尽可能多的"语义框架"（semantic frames）、"句式框架"和认知域，并建立相应的各种模框（frames）和图式（schemas）。

5）各种排歧研究。包括各种排歧类型和针对不同歧义类型所采取的种种不同的排歧策略。

6）在上述研究的基础上，建设汉语知识库，特别是明确标有词的语义特征和词的句法特征的词库（可分别建库），同时开展跟句处理相关的实验和应用系统的研制开发。

## 二、歧义短语及其排歧策略

### 1. 歧义类型

自然语言中充满了歧义，这种现象在字处理、词处理层次上都有反映，

如同音字问题、歧义字段与自动分词问题、多类词的词性确定问题等。随着中文信息处理向更深层次的推进，自然语言中的歧义现象及相关的排歧策略越来越制约着中文信息处理的进程。句处理中的歧义问题表现得更为复杂，处理的难度也更大，如短语边界的确定、词语串合语法性的判定、歧义短语的分化、与实词有关的语义问题、成分间的语义关系问题等等。

歧义是自然语言的计算机理解中所面临的一个最严重的问题。人依靠大量的背景知识和使用母语所积累的经验，可以消除日常语言交际中的大部分歧义，产生歧义理解的可能性不大。但是计算机不可能把一个人的全部知识都储存在机器中，所容纳的与语言理解有关的背景知识与人相比更是少得多，产生歧义理解的可能性因而会大大增加，这种情况在汉语中特别严重。因此分析汉语中歧义产生的各种原因，并以此建立起一些有效的规则，就成了语言学家最为关心的问题之一。

自然语言的歧义问题，实质上是意义与形式之间的矛盾问题。同一形式与不同的意义相联系，必然会产生歧义，这是自然语言不同于人工语言的一个主要方面。

(1) 词汇歧义和结构歧义

根据歧义发生的语法单位层级，可以将歧义分为词汇歧义和结构歧义两类。词汇歧义主要体现在多义词和兼类词上，它对文本的自动切分以及自动词性标注有很大的影响。

结构歧义发生在词与词的组合中，这一问题早就引起了学者们的注意，总结出不同类型的歧义格式。如：

V＋N
学习文件
参考资料
V＋的＋是＋N
关心的是母亲
反对的是校长
P＋N＋的＋N
关于鲁迅的书
对校长的意见

歧义格式的发现和归纳对于汉语歧义现象的研究具有重要意义，为歧义

的分化提供了可能。但是由于语言中的任何一个结构歧义的形式都不是孤零零地存在着的，它往往代表具有某种格式的多种实例，而满足这种构成模式的实例并不总是有歧义。如同样是“V＋N”的格式，“学习汉语”与“学习成绩”并没有歧义。因此歧义格式的发现只是说明某种格式有歧义的可能性。

(2) 潜在歧义

冯志伟在研究计算机处理歧义问题时，提出了“潜在歧义”的理论，认为“歧义格式”所反映的类别的歧义在具体的语言中有时存在，有时不存在。因此歧义格式所反映的歧义并不是现实的歧义。当用具体的词语去代替歧义格式中的范畴符号时，在所形成的具体的短语或句子中，潜在歧义有可能继续保持，也有可能消失。可见，抽象的歧义格式所包含的歧义实际上只是一种潜在的歧义。这是自然语言歧义格式研究与一般的句法上歧义研究的不同之处，是自然语言歧义格式研究的一个主要特点。

用潜在歧义理论来分析汉语的歧义结构时，可以得到比较满意的结果。例如汉语的“VP＋的＋是＋NP”是一个歧义格式，如“反对的是少数人”有歧义，既可以理解为“提反对意见的是少数人”，也可以理解为“所反对的是少数人”；而同类格式的“关心的是分数”，则没有歧义。

(3) 包含终结符的歧义格式和不包含终结符的歧义格式

根据歧义格式中是否包含特征词(终结符)，可以把歧义格式分为包含终结符的歧义格式和不包含终结符的歧义格式。如：

一张电影院的海报
一家电影院的经理
捐赠的时间和地点
倒塌的房屋和难民

上两组实例可分别表示为“MP(数量成分)＋NP(名词性成分)＋的＋NP”和“VP＋的＋NP＋和＋NP”，这两种格式分别有两种组合方式，都是歧义格式：

MP＋NP＋的＋NP
___ __________(一张电影院的海报)
__________ ___(一家电影院的经理)
VP＋的＋NP＋和＋NP
_____ ___________(捐赠的时间和地点)
____________ ______(倒塌的房屋和难民)

两种歧义格式中包含助词“的”和连词“和”，属于包含终结符的歧义格式。终结符特别是虚词，往往与特定的句法结构相关，可以视为某种结构关系的标记，如“的”可作为定中结构的标记，“和”可作为联合结构的标记，终结符对歧义结构的分化有重要的参考作用。

歧义格式中不包含特征词的属于不含终结符的歧义格式。如：

公司项目经理
羊皮领子大衣
老师辅导学生
电器修理教材

两组格式可表示为：NP＋NP＋NP、NP＋VP＋NP，它们也都存在内部结构歧义：

NP＋NP＋NP
____ ________（公司项目经理）
________ ____（羊皮领子大衣）
NP＋VP＋NP
____ ________（老师辅导学生）
________ ____（电器修理教材）

（4）外显型歧义格式和内含型歧义格式

根据一个格式所代表的不同结构对外部环境的影响或受外部环境的制约情况，可以将歧义格式分为外显型歧义格式和内含型歧义格式。

外显型歧义格式所代表的不同结构具有不同的功能类型，因而对外部环境的影响或受外部环境的制约较大。如：

VP＋NP＋的＋NP
修理自行车的后胎
修理自行车的老王

该歧义所代表的实例有的是动宾结构，可以作谓语；有的是定中型偏正结构，不能作谓语。不同的组合反映出整体功能类型的差别，也直接制约着它们的句法功能。

内含型歧义格式所代表的实例整体功能类型没有明显差别。如：

VP＋NP＋NP

大钢铁公司

大眼睛姑娘

该歧义格式所代表的实例都是定中型偏正关系，功能类型都是名词性的。

区分外显型与内含型两种歧义格式有助于正确地考虑排歧策略。

(5) 真歧义格式、准歧义格式和伪歧义格式

根据某种歧义格式形成事实歧义的可能性，可以将歧义格式分为真歧义格式、准歧义格式和伪歧义格式3类。

真歧义格式很容易在自然语言中找到对应的实例，无论是对人还是对计算机都需要采取相关的排歧策略。准歧义格式对人来讲可能不造成理解上的困难，但对计算机而言，却涉及到排歧问题。伪歧义格式虽然可以有不同的组合方式，但无论对人还是对计算机，不同的组合方式并不造成理解上的困难。如"踢破球"所代表的"VP＋AP＋NP"是一个真歧义格式。"被警察抓住罚了款"和"被政府邀请参加庆典"是"PP＋VP＋NP"中的两个实例，但对人来讲，理解起来并无太大困难，但计算机理解时却涉及到内部结构定界的问题，将它们分别定界为："[被警察[抓住罚了款]]"和"[[被政府邀请]参加庆典]"。"DP＋VP＋NP"所代表的实例虽然可以有两种不同的组合方式，但无论按照哪一种方式，对理解的影响甚微，属于伪歧义格式，如"[认真[学习汉语]]"和"[[认真学习]汉语]"。

包含终结符的歧义和不包含终结符的歧义、外显型歧义和内含型歧义以及真歧义、准歧义和伪歧义之间的区分有交叉，是在对潜在歧义讨论的基础上对信息处理中的排歧问题的进一步深入分析，有助于针对不同类型的歧义寻找相应的排歧策略。

**2. 歧义的消歧方法**

(1) "制约"和"优选"的排歧方法

"制约"方法和"优选"方法是自然语言的计算机处理中普遍采用的两种排歧方法。

所谓"制约"，就是利用句法、语义条件，排除不能满足制约条件的结构，从而达到消除歧义的目的。如：3个学校的实验员，这种歧义是由量词"个"造成的。可以通过量词与名词之间的选择限制关系来分化歧义。如将造成歧义的量词"个"换为"所"和"位"则歧义消除，"3所学校的实验员"和"3位学校的实验员"都是单义的。

所谓“优选”，就是在若干个存在歧义的候补结构中，选出一个最优的结构，从而达到消除歧义的目的。如“小王说故事很有趣”可以有不同的理解，但根据述宾结构的优先度大于主谓短语作主语的结构的原则，排除以主谓短语作主语的结构，消除歧义。但是这种方法“优先度”的确定根据的是语感，而语感往往带有强烈的主观色彩，因此缺乏客观性，这种方法通常要和其他方法一起运用。

(2)“受限汉语”方法

由于“优选”方法有一定的局限性，语言学家在自然语言理解中采用的消歧方法大都是“制约”。北京大学俞士汶提出了“受限语言”的理论，认为“实用的自然语言处理系统对自然语言总是自觉不自觉地进行了某些限制，这使人们认识到受限语言的研究是有价值的”①。受限汉语就是对汉语加以限制，造成一种带有人为色彩的自然语言，降低自然语言的复杂度，在语言系统的精确度和表达意义的自由度之间找到一个合适的平衡点。例如从歧义格式和实例歧义间存在着对应关系，歧义结构对上下文有不同的影响程度等方面出发，在对汉语的短语结构进行受限研究时，就要考虑上述因素，可以选择无歧义或歧义程序较低的结构作为汉语句法结构系统最基本的核心部分，然后在此基础上，考虑表达需要作适当的增补，就能得到比较合理的受限汉语的短语结构系统。

## 三、句处理中的语义问题

### 1.“语义”内容的复杂性

语义问题是自然语言信息处理特别是句处理阶段的重要环节。在自然语言的语义分析或语义知识表达方法方面，大致已经有语义场理论、义素分析理论、配价理论、格语法、论旨理论、概念依存理论、语义网络、蒙太格语法，等等。就大规模描述自然语言语义知识的工程实践而言，以英语和其他主要欧洲语言为描述对象的研究，影响较大的有 WordNet、MindNet、FrameNet 等等。中文信息处理方面则有“信息处理用现代汉语语义分类体系”、“现代汉语述语动词机器词典”、“董氏语义知识词典”等比较有代表性的研究工作。语义知识在自然语言处理中的作用可以概括为两个主要的方面：① 帮助句法分析得到正确的结果；② 为发现句子中各个语言成分

---

① 俞士汶.语法知识在语言信息处理研究中的作用.语言文字应用.1997,(4)

所对应的概念之间的意义关系提供支持。

句处理中所要考虑的语义是多方面的，且分为不同的层次：如句子的情态(modality)义和命题(proposition)义，句子本身的命题义与句子在使用环境下可能具有的语用义，句子本身的命题义中，还包括由实词带来的意义和由虚词带来的意义，还有句子中某个单元(包括词或短语)本身的意义以及单元与单元组合所产生的种种意义，所有这些意义要分层处理。此外，由于句子中存在着成分的移位、增添与省略等现象，句处理还要解决好句中缺省部分的添补与理解的问题。

2. 实词的自身义

就实词而言，存在着多方面与句处理有关的语义问题。从实词本身来说，则涉及到实词的概念义、指称义和语义特征 3 种。

(1) 概念义

词的概念义分为外延义与内涵义两种。从概念外延的角度所理解的概念义，就是概念的外延义；从内涵角度所理解的概念义，就是概念的内涵义。如：

这两位农民是从四川来的。

王教授在农村劳动了一年，像个农民了。

前一例中的"农民"表达的是外延义，后一例中则表达的是内涵义。外延义含明显的指称性，是指所有具有该词内涵义特征者。

(2) 指称义

指称义包括有指与无指、通指与专指、定指与不定指。

有指即某个名词所指确有具体的人或物，无指即名词所指没有具体的人或物。如：他考过研究生。该句有歧义，"研究生"既可是指人名词，前可以加表物量或动量的数量词语，如"考过 3 回/3 个研究生"；也可以指标示学业程度的抽象名词，前只能加表动量的数量词语，但不能加表物量的数量词语。

通指即句中名词表示的事物是一个类名，专指即句中名词表示的是事物的个体。如：

| | |
|---|---|
| 他不吃鱼。 | *他不吃 3 条鱼。 |
| 我吃过鱼了。 | 我吃过 3 条鱼了。 |

由于通指成分表达类的概念，因而不能受表物量的数量词语的修饰，而专指成分表达个体概念，可以受表物量的数量词语的修饰。

定指即说话人认为听话人知道名词所指称的对象，不定指即说话人认

为听话人不知道名词所指称的对象。如：

客人来了。　　　　来客人了。
我看过那本书。　　那本书我看过。
我借了两本书。　　*两本书我借了。

定指成分出现在主语位置上是自由的，而不定指成分在主语位置上出现受到很大限制。

(3) 语义特征

实词的语义特征对句法结构同样有制约作用，如与动词“给”相关的句式中动词的语义特征的不同会给句法结构带来影响。再如动词的“自主”与“非自主”的性质也会影响动词的句法性质。如：

看！　　去看去！　　别看！　　别看了！
*忘！　　*去忘去！　　*别忘！　　别忘了！
*醒！　　*去醒去！　　*别醒！　　*别醒了！

含“自主义”特征的动词可以构成肯定式或否定式祈使句，而带有“非自主义”特征的动词构成祈使句会受到很大的限制，有的只能构成否定式祈使句，有的则不能构成祈使句。

动词的持续与非持续等语义特征也会给句法结构以及语义表达带来影响。如：

看了3天了　　等了3天了
挂了3天了　　腌了3天了
死了3天了　　断了3天了

“看、等”等动词所在的结构只表示动作行为持续的时间，“挂、腌”等动词所在的结构既可表示动作行为持续的时间，也可以表示某种状态持续的时间；“死、断”等瞬时动词所在的结构只表示某种状态完成后经历的时间，这种差别是由动词自身的语义特征决定的。

实词的自身义能对句法结构产生影响还表现在歧义结构中。如现代汉语中“N＋A＋的＋不＋V”的词类排列模式代表了两种不同的结构关系：

N＋A＋的＋不＋V
主　　谓　　（甲）
　　主　　谓　（乙）

两种结构都集合着大批的实例。如：

① 香蕉青的不买。

② 皮儿青的不买。

③ 皮儿青的不吃。

④ 张三青的不吃。

以上四例的表义情况不同，有的是单义的，有的是歧义的。从形式上看它们彼此间都有共同的词语，如①和②之间的差别是句首的名词不同，①中的动词“买”可以支配“香蕉”，但“香蕉青的”在实际生活中不能构成用于指称的“的”字短语，因而是单义的，只能按(甲)切分。②句首名词“皮儿”不能充当动词“买”的受事，而“皮儿青的”可以用于指称并成为“买”的受事，所以只能作乙切分。②和③的差别在于所用的动词不同。③动词“吃”既可以以句首名词“皮儿”为受事，也可以以“皮儿青的”这一指称性短语为受事，因而是歧义的，可以有两种不同的切分。③和④的差别在于句首所用的名词不同，前者是指物的“香蕉”，后者是指人的“张三”。两者不仅构造层次不同，④只有(甲)一种切分可能性；而且即使在相同的构造层次下，它们的语义格式仍有差别。③按照(甲)切分时，成分间的语义关系是“受事[整体]||受事[部分] 动作”，④的语义格式为“施事||受事 动作”。

这种现象超出了以往关于汉语歧义现象的观点所能解释的范围，表明句中词语的具体意义对所在结构是不是歧义格式，有没有歧义具有重要影响。

3. 实词之间的关系义

实词之间的关系义指实词与实词相互关系中所呈现的意义，词与词的组合一方面可以形成一定的句法结构关系，另一方面彼此间还存在着不受句法结构关系影响的语义关系。实词之间的关系义包括名词的格范畴、论元、论旨角色、配价，词与词之间语义上的制约关系，语义指向等等。

(1) 配价语法、格语法

配价理论的创始人是法国语言学家吕西安·特思尼耶尔，他从化学中借用“配价”的概念是为了反映语言结构中名词性成分和动词性成分之间的一种最基本的依存关系。该理论认为句子是一个有组织的整体，句子的构成成分不只是表面上看到的词，更重要的是词与词之间的关联，动词好比一个带钩的原子，用来钩住一定数目的名词性成分并使之处于依附状态。动

词所带钩子的数目就是该动词所能支配的名词性成分的数目，即动词的“价”（或称为“向”），据此可以将动词分为零价动词（无须支配必有名词性成分的动词，如英语中的“rain”）、1价动词（支配一个必有名词性成分的动词，如“游泳”、“走”）、2价动词（支配两个必有名词性成分的动词，如“看”、“知道”）和3价动词（支配3个必有名词性成分的动词，如“给”、“告诉”）。

配价语法的主要目的在于揭示动词对名词性成分的支配能力，其基本观点是动词中心论，认为动词是句子的结构核心，在配价语法中，动词具有特殊的重要性，通过对动词配价能力的描写，可以简化对句子中名词性成分与动词性成分之间依存关系的说明，也可以有效地预测句子可能具有的依存结构，对句子的结构类型作出较为合理的划分。这种理论在20世纪70年代引入我国，并被广泛运用于与动词相关的语法研究和语言信息处理领域，总结出了含动词的“的”字短语表义的可能性，即歧义指数，并将这种方法进一步扩展到形容词、名词的研究，提出了形容词的配价和名词配价的概念，增强了语言研究结构的可计算性，在研究成果的形式化方面表现出了较大的优势。

配价理论旨在揭示动词对名词性成分的支配能力，主要从语义的决定性与句法的强制性的角度研究与名词、动词、形容词的意义有关的语义成分。要全面地反映有价词语的支配能力，还应该描写这些词语与受其支配的从属成分之间的语义关系，即从属成分的语义角色。这方面正是格语法的研究重点。

格语法是20世纪60年代由美国语言学家菲尔墨提出的，在其初期理论中，将句子的基础结构分为情态和命题两部分，命题由动词和一个或几个名词短语组成，每个名词性成分与动词都具有一定的语义关系，即格关系，如施事、受事、与事、工具、结果、处所、时间等。

不同的动词要求语义格的能力是不同的，而且这些语义格在句法结构中有一定的出现位置，据此可以描述以动词为核心的语义配置式。借助于不同的语义配置式可以简明地表达一种语言中的各种语义结构，并且能反映不同的语义格在句法结构中的表现情况，从而为说明句法结构与语义结构之间的关系提供了便利。因此这种理论提出后，很快便引起了自然语言处理领域的极大兴趣。如我国学者鲁川、林杏光等以格语法为基础研究面向中文信息处理的现代汉语语义格关系系统，设计出两大类七小类共22个角色的现代汉语语义格系统：主体格包括施事、当事、领事，客体格包括受

事、客事、结果，邻体格包括与事、同事、基准，系体格包括系事、分事、数量；情景中的凭借包括工具、材料和方式，环境包括范围、时间、处所和方向，根由包括依据、原因和目的。据此将现代汉语的动词分为他动词、自动词、外动词、内动词、领属动词和系属动词 6 类，并对 1 000 多个动词的语义配置式包括基本式和扩展式作了描写。虽然汉语动词的语义格关系的描述是没有止境的，诸如动词与名词之间语义格关系的数量、类型，动词与语义角色的句法配置式等，都还有大量的工作要做，但这项研究为解决中文信息处理中的语义关系问题探索出了一条新路。

(2) 广义配价模式

配价语法和格语法虽然研究的侧重点不同，但它们却具有不少的共同点：都坚持动词中心论，都重在揭示动词与名词之间的语义关系，"'名词格范畴'也好，'动词论元'也好，其实说的是一回事儿，都是指一个动词所表示的行为动作在进行时必然或可能会关涉到的某种范畴的事物，只是角度不同而已"①。

关于语言成分的搭配能力的知识对正确分析自然语言的句子具有关键作用，一个理想语言知识描述系统，应该把语言中任意两个成分间的搭配可能性纳入描述范围之内。而以往的语义研究侧重在动词与名词的语义搭配方面。但是，词语之间的搭配还应包括名词与名词、名词与形容词、动词与形容词等成分间的语义搭配性质。在描写实词的语义搭配信息之外，还有必要描写短语的语义搭配信息，因为在从词到短语的组合过程中，语言成分的搭配性质会发生各种变化，记录这些变化对分析是有作用的。詹卫东(2000)在已有的自然语言语义分析理论基础上，结合汉语信息处理的实践，提出了一个增强的汉语语义知识表达框架：广义配价模式，尝试在 3 个层级上描述汉语实词的语义知识，把对实词的语义信息描述扩展到对短语的语义搭配信息的描述。

在广义配价模式下，汉语"名、动、形"三大类主要实词的语义知识可以在下面这 3 个层级上展开描述。

1) 词语本身所属的语义类。在词库中，需要对每一个具体的实词描述其所属的语义类属性。比如名词"太阳"语义上属于[天体]类、形容词"红"

① 陆俭明.关于句处理中所要考虑的语义问题.汉语语法研究的新拓展(一).浙江教育出版社，2002.11

语义上属于[颜色]类、动词“拿”语义上属于[搬移]类,等等。跟对词进行句法功能分类的目的相似,对词语进行语义分类实际上也可看作是在刻画一个词的搭配能力。比如,就名词而言,同属一个语义类的名词一般可以搭配形成无标记联合结构,如“太阳月亮”;不同语义类的名词一般不大能够搭配形成无标记联合结构,而是倾向于形成偏正结构,如“阳光城市”。此外,语义类属性的另一个作用是使得搭配可以在词对类的概括水平上进行描述。比如,动词“支持”可以搭配许多“个别”的名词,如“支持现任总统”、“支持张三”、“支持这个计划”等等。如果在更为概括的语义类层面上加以描述,就不需要一一罗列能跟“支持”搭配的对象,而是通过把这些个别名词概括为语义上属于“人类”、“抽象事物”等的办法来加以描述。不同语义类别的词语在句法上应该有显著的差异表现,而同属一个语义类的词语在句法上应该有显著的相似表现。跟词类的情形一样,要达到这样的目标光靠分类是很困难的,对此,可以在更多维度上对一个词的语义性质进行区分。

2)词语配价成分的个数、角色类型及其选择限制。这个层级是更直接地来刻画一个实义词跟其他实义词的搭配可能性。从理论上讲,词语的配价成分这个概念对动词、形容词、名词都适用,不过从重要性上讲,动词更为突出,动词的配价成分的情况也最为复杂。

对动词来讲,配价数(即配价成分的个数)的取值范围为:0、1、2、3。角色类型分成两部分:一部分是核心语义角色,包括主体、客体、与事、工具、处所等,“主体”大致包括一般所说的施事、当事等语义格,“客体”大致包括一般所说的受事、对象、目的、结果等语义格;一部分是外围语义角色,包括空间和时间等。对配价成分的选择限制则要求指明一个动词对其配价成分有哪些条件限制(主要是从语义方面来进行限制)。限制可以在句法层面描述,也可以在语义层面描述;可以是精确限制,也可以是概括限制;可以从正面来要求配价成分应符合某种条件,也可以从反面来要求配价成分应避免哪些情况。表达形式可以根据具体动词的不同情况灵活选择。如果强调精确度,就最好是词对词的搭配。如果要强调概括度,就描写词对类或者类对类的搭配。比如“后胎”是1价构件类名词,其“主体”名词限于“自行车”、“三轮车”、“汽车”等等带“车”字的词语。要描写“后胎”的配价性质,就可以强调精确度,用“字”特征来描述,而不用语义类特征来描述。而动词“吃”能搭配的名词很广泛,要描写它的配价性质,就可以概括描写为跟[可食物]类名词搭配。

3）词语配价成分的变化情况。所谓的词语配价成分的变化情况，主要是指动词配价成分的变化情况。引入这个层级的语义信息，主要目的是描述动词跟补语成分（通常是形容词或动词）的搭配能力。比如“走”的一个义项是1价位移类动词，它可以搭配一个名词性成分，即“主体”，并且对“主体”的选择要求是“人”或“动物”等可移动的物体。如果考虑“走”的“主体”配价成分的变化情况，不难发现，经历了“走”这个过程，“主体”可能发生的变化是位置变化以及性状变化，述补结构“走远了”、“走近了”正是表示“主体”的位移变化；而“走累了”、“走跛了”则是表示“主体”的性状变化。“走”体现另一个义项“离开”时，配价成分的变化情况跟表示位移的“走”的情况有显著的不同，它的“主体”没有“远”、“近”等位移变化。这并不是说主体在物理意义上没有移动，而是指在说汉语者的心理认知中并不关注位移，这时关注的是“在场”还是“缺席”，这个义项下的“走”可以出现在“他已经走掉了”中，“走”跟补语动词“掉”搭配。表示“离开”的“走”，其“主体”配价成分也没有“累”、“跛”等性状变化，但这时“走”可以跟形容词“光”搭配，即有“主体”的数量变化。比如“人都走光了”、“他们中已经走了3个”。人之所以能够准确判断出“走”的不同意义，正是依靠了不同的“走”所出现的不同上下文，或者说就是不同的搭配。在广义配价模式的配价成分的变化情况这个层面上，表示“位移”义的“走”的配价成分的变化情况可描述为[主体变化：性状|位移]；表示“离开”义的“走”的配价成分变化情况可描述为[主体变化：数量]。

自然语言中动词配价成分的变化是普遍存在的，这些变化在句法上大多是由汉语的述补结构来表达的。通过描述动词配价成分的变化情况，有可能为分析汉语的述补结构，特别是其中补语成分的语义指向提供一条线索。

广义配价模式还特别强调在短语一级描述语言成分的语义信息。因为在从词向短语的组合过程中，一个成分的语义搭配能力可能发生变化，会导致搭配成分范围的扩大或缩小。描述这种变化，有助于计算机得到正确的分析结果。比如，动词“带”一般只能搭配表示具体物体的名词，如“带钢笔”、“带钱”等。但带上补语“来”之后，其搭配范围就扩大了，可以说“带来了一线希望”、“带来了一个好消息”等。这是语义搭配能力扩大的情况。再比如动词“穿”带上介词补语“在”构成动词性短语“穿在”，不能搭配普通名词成分，只能搭配处所成分。比如既可以说“穿西装”，也可

以说“穿身上”；但只能说“穿在身上”，不能说“穿在西装”。这是语义搭配能力缩小的情况。在广义配价模式下，对短语“带来”、“穿在”等的这些语义搭配性质的变化进行描述，将有助于正确分析像“穿在身上的衣服”这样的短语。它的结构应该是：[[穿在身上的] 衣服]，而不是：[[穿 在][身上的 衣服]]。

广义配价模式最基本的思想是分层次描写实词的语义性质，这跟以词类信息加属性特征描述来刻画词语的语法性质的思想是一致的。

(3) 语义指向分析

所谓语义指向，指的是句中某个句法成分与另一成分之间存在的语义联系。如：

他圆圆地画了一个圈。
他愉快地画了一个圈。
他慢慢地画了一个圈。

以上 3 句的句法分析结果完全相同，但状语语义上所联系的成分不同，“圆圆地”指“圈”，“愉快”指“他”，“慢慢地”指“画”这一动作，也就是说以上 3 句状语的语义指向不同。

确定句中成分与其他成分之间的语义联系是语言理解的一个方面，也是自然语言处理中所要考虑的一个问题，语义指向的分析可以从语义联系方面对某种语法现象作出说明和解释。

从理论上说，应该是每个句法成分都有语义指向的问题，但并非所有的句法成分都需要通过语义指向的分析来加以说明。语义指向分析的对象主要是句法组合中的补语、修饰语以及谓语这 3 种成分。

补语是述补结构中补充说明中心语的成分，可以由动词或形容词来充当，而动词、形容词在语义上要求支配不同数量的语义成分，因此出现在补语位置上的谓词都与其他成分有语义联系。如：

踢坏了（“坏”语义上指向“踢”的受事，如“球”等）
踢赢/累了（“赢/累”语义上指向“踢”的主体，如“他们”等）

再如：

他分别向会长和秘书长汇报了工作进展情况。
他和老王分别向会长汇报了工作进展情况。

例中的“分别”充当句法上的状语，但在语义上联系的成分不同。前一句中“分别”在语义上指向“汇报”的对象：“会长”和“秘书长”，后一句中“分别”在语义上指向“汇报”的主体，即“他”和“老王”。可见，副词“分别”充当状语时，语义上可以前指也可以后指，如果语义指向不能惟一地确定，则句子有歧义。如：

他和老王分别向会长和秘书长汇报了工作进展情况。

句中“分别”既可指向“他和老王”，也可以指向“会长和秘书长”，从而造成了四重歧义。

语义指向分析对说明句法成分间的语义联系以及语义关系对句法结构的制约性具有重要作用。如：

究竟他去了哪里？
他究竟去了哪里？
究竟谁去了网吧？
* 谁究竟去了网吧？

可见，副词“究竟”的出现位置有一定的限制，这是因为它的语义指向特点所致。“究竟”在语义上只能指向疑问成分，如疑问代词，肯定否定并列的形式等；在方向上，它只能后指，不能前指，即它所联系的语义成分在线性序列中只能在它的后边出现，否则就是不合法的结构，如上带“ * ”的例句。

4. 句处理中涉及的其他语义问题

句处理中的语义问题还涉及到特定的范畴义（某种特定的范畴所赋予的意义）、实词之间的组合义（句法结构关系所赋予的意义以及语义结构关系所赋予的意义）、句式义（特定的句式所赋予的意义）、语用义（语用因素所赋予的意义）以及认知义（认知因素所赋予的意义）等问题。

范畴义包括数量范畴、领属范畴等，会对句法结构产生影响。现代汉语中，数量范畴是由数量短语来表示的，包括：数词＋量词、代词＋数词＋量词等。数量范畴对汉语句法结构的制约作用表现在：

有些句法组合强制性地要求有数量成分。如：

| | |
|---|---|
| 急了他一身汗 | * 急了他汗 |
| 买了小店一只球 | * 买了小店球 |
| 大大一个西瓜 | * 大大西瓜 |

怎么一种脾气　　　　　　＊怎么脾气

有些句法组合如果没有数量成分，则只能是粘着型结构。如：

吃了一个苹果。　　　　　？吃了苹果。
他送我一个手机。　　　　？他送我手机。

有些句法组合排斥数量短语。如：

什么衣服？　　　　　　　＊什么两件衣服？
什么书？　　　　　　　　＊什么三本书？

实词性词语彼此组合后所产生的意义，是实词之间的组合义，包括句法结构关系所赋予的意义和语义结构关系所赋予的意义。如“学习材料”、“进口设备”具有述宾和偏正两种不同的结构关系，分别赋予它们谓词性和体词性的功能，作为谓词性结构，具有陈述功能，可以充当谓语，而作为体词性结构不能充当谓语。再如“鸡不吃了”中“鸡”与“吃”之间可以是施事和动作的关系，也可以是受事与动作的关系，因而是歧义的。而在“鸡不吃食了”和“鸡他不吃了”中“鸡”与“吃”之间的语义关系能惟一地被确定，因而是单义的。这是语义结构关系所赋予的意义。

与不同的句型、句式或句类有关的语义便是句式义。如：

① 他在墙上贴画儿呢。
② 画儿在墙上贴着呢。
③ 画儿贴在墙上呢。

①是动态句，表示“贴”这一行为正在进行，②③是静态句，“贴”不表示动作行为，而表示“画儿”存在的方式，区别在于②句强调存在的方式，而③句强调存在的处所。这种动态与静态、存在方式与存在处所的差异是由特定的句式带来的。

句子是动态的交际单位，句子层面上还存在着一些在短语中不存在的、与语言交际密切相关的因素。由语用因素带来的语用义以及由语言形式以外的认知因素带来的认知义则是句处理所面临的更为复杂的语义问题。

## 参考书目

1. 陈力为、袁琦主编. 中文信息处理应用平台工程. 电子工业出版社，

1995

2. 冯志伟.论歧义结构的潜在性.中文信息处理.1995(4)

3. 冯志伟.中文信息处理与汉语研究.商务印书馆,1992

4. 冯志伟.自然语言的计算机处理.上海外语教育出版社,1996

5. 黄昌宁、李涓子.语料库语言学.商务印书馆,2002

6. 陆俭明.现代汉语语法研究教程.北京大学出版社,2003

7. 俞士汶.关于受限汉语的规则汉语的设想.语文现代化.山东教育出版社,1995

8. 袁毓林.语言的认知研究和计算分析.北京大学出版社,1998

9. 詹卫东.面向中文信息处理的现代汉语短语结构规则研究.清华大学出版社、广西科学技术出版社,2000

10. 周明等.统计与规则并举的汉语句法分析模型.计算机研究与发展,1994(2)

# 第五章　语言规划与语言调查

## 第一节　语 言 规 划

### 一、语言规划的定义、性质和原则

1. 什么是语言规划

语言规划(language planning)又叫语言计划，是政府或社会团体为了解决在语言交际中出现的问题，有计划、有组织地对语言文字进行的各种活动和工作的统称。它是一个国家语言政策的具体表现。

语言规划是 1957 年由语言学家威因里希(Ureil Weinrich)首先提出来的。其实在这一名词出现之前，语言规划早已实行了。例如我国秦代，秦始皇就把文字统一为小篆，实行“书同文”的政策，这就是历史上一次较大规模的语言规划；欧洲罗马帝国时，也实行了以拉丁文统治多民族的政策。文艺复兴时期，意大利、法国、西班牙等国家都成立过负责语言的机构，以保持和维护民族语言的纯洁，促进语言的交际功能，为工业化革命奠定了文化基础。这也是较早的语言规划。随着社会的发展，语言规划也在发展，目前已发展成为应用语言学的一个重要的组成部分。

语言规划需要应用研究和基础理论研究的结合。不仅要解决社会使用的实际问题，还要充分利用语言自身发展的规律在实践中不断检验、修改和丰富我们的语言，做到相互验证、相互推动。语言规划应该面向社会，它既不应该是对语言现象的客观描写，也不应该是对语言的主观限定。

2. 语言规划的性质

根据语言规划的定义，我们认为它有以下一些性质：

1）语言规划是政府或社会团体行为，一般不可能是个人行为。语言规

划又叫语言计划，这种计划通常是由政府或政府某一部门制定，通过这些计划，政府的语言政策才可以得到确立和执行。它体现了政府或社会团体对社会语言的根本态度，同时也表现出语言规划的社会性。虽然有时著名人士或作家等会对语言规划起到一定的影响，但这种影响只是语言运用中的，不会影响语言的发展和方向。

2）语言规划的目的在于解决在语言交际中出现的问题，以促进语言的社会交际功能的发挥。这也就是说，语言规划的目的是由语言的社会交际功能决定的。语言规划就是立足现在、面向未来，从正面引导和规范语言的运用。因此，离开了语言的交际功能就无所谓语言规划，就会变成纸上谈兵。同时我们也要注意，语言规划带有一定的民族情绪，如果这种民族情绪影响了语言交际功能的发挥，就可能导致语言规划的失败。这是语言规划交际目的性的表现。

3）语言规划是有计划、有组织、有步骤地进行的，它往往要持续很长时间，不可能一蹴而就。从某种意义上来说，一种语言就是一个民族甚至国家的标志，它往往会涉及社会各方面的问题，甚至受到政治的影响，所以语言规划有时需要几十年、要经历非常曲折而艰苦的过程才能取得成功，必要的时候甚至要借助于法律。这也能够充分表现出语言规划的长期性。

4）语言规划的内容庞杂，除了对语言文字一般性的管理之外，还包括语言的选择、语言的标准化和规范化、文字的创新、文字的改革等等。同时语言的规划还要涉及多种语言和不同的专业学科，从事语言规划工作不仅需要广博的语言知识，还需要有丰富的社会经验和工作能力。

5）语言规划一般是由知名人士或政府领导提出的，他们的社会知名度就是语言规划权威性的所在。例如我国的文字改革工作一直得到毛泽东、周恩来等国家领导人的关心和支持，这使得文字改革工作有了相当雄厚的社会基础，这也是文字改革工作为什么能够取得如此巨大成就的原因之一。权威所带来的社会声望对语言规划的成败起着重要的作用，语言规划机构本身也需要这种社会声望，这种社会声望不仅需要规划机构展示出高超的学术水平，也需要政府和领导的支持。一个社会声望不好、没有权威性的语言规划机构是不可能搞好语言规划的，这说明语言规划要有权威性。

**3. 制定语言规划的原则**

(1) 科学性

语言是人们最重要的交际工具，是一个相对稳定的符号系统，但它又在

不停地发展变化。

语言的发展是有规律的，我们要根据语言的发展规律引导语言的发展。社会现实也会影响语言的发展，也是有规律的。这些规律都值得我们认真学习、认真研究，它们也都是语言研究的重要内容。

语言规划需要把理论和实践结合起来，既要解决语言的应用问题，也要解决语言理论问题，在制定语言规划过程中，要充分注意和利用语言本身的发展规律引导、丰富和完善语言，使理论和实践有机地结合起来共同推动语言的健康发展。制定语言规划要充分发挥专家的作用，专家有丰富的专业知识，而且大多是关心、了解社会的，对社会也负有责任心。对不同问题进行讨论，有助于正确的语言规划的制定，而且制定之后也许要一边实行，一边修改，不断地完善。当然语言规划的制定和修改都需要慎重，需要经过必要的程序，要以科学为根据。

(2) 适用性

确定语言某方面的标准是指令性的，这就要求必须简明，标准尽量做到整齐划一。所以提出的意见也要具有一定的倾向性，有一个总的方向，要做到能够吸引人们主动靠拢，在实践中不断地加以讨论完善。语言是不断发展变化的，但笼统而言，语音、词汇、语法的活跃程度是不同的；相对而言，语法规则是最稳固的，口语和文学语言相当活跃，新词语也特别活跃，最活跃的是许多临时的修辞性用法。而且文化层次不同的使用者在使用语言时也有一些自己的特点，这都需要我们加以区别对待。

在用字方面，指令性的多；在读音方面，指令性的多，但也需要有一定的指导性；在用词和语法上，指导性要多于指令性。对于具有中等文化程度的人来说，指令性要强于指导性，而对于中等文化程度以上，尤其是具有高等文化程度的人来说，要更多地运用指导性。所以规划的制定要根据具体情况，要有可操作性。

(3) 稳妥性

语言规划的制定要顺应社会和时代的发展，要因势利导，符合语言发展的规律；要顺乎大多数人民群众希望语言便于交际的要求，把握好尺度，做到既不守旧，也不冒进；对语言中的一些问题要循序渐进，等到时机成熟了才可以加快步伐；而且语言规划要保持一定的延续性，不能朝令夕改。

(4) 动态性

语言是不断发展变化的，使用语言的人是不断发展变化的，社会也是不

断发展变化的，这就要求我们的语言规划也要能够随着社会的发展变化而变化。语言规划的制定和实施都需要有一定的预见性。指令性和指导性的标准都需要不断地修改和完善。譬如，普通话水平测试的标准，要随当地大部分人的普通话水平的提高而提高，甚至必要的时候，总体的工作方针也可以改进。

## 二、语言文字的地位规划和本体规划

语言规划从总体上讲，可以分为地位规划和本体规划两个方面。

### 1. 语言文字的地位规划

语言文字的地位规划，就是决定某种语言或文字在社会交际中的地位，这种规划，一般要借助于政治的力量来进行。

第二次世界大战之后，印度、新加坡、菲律宾、马来西亚、东非和拉丁美洲等国家和地区都进行了共同语文的建设，而且也都取得了很大的成绩。

新加坡 1959 年成为英国的自治邦，1963 年脱离英国，加入马来西亚联邦。新加坡是一个多语言的地区，主要讲英语、汉语、马来语和泰米尔语，但是马来西亚联邦只能以马来语作为官方语言，不容许新加坡保留其他语言作为官方语言，这样两者产生了激烈的矛盾。新加坡于 1965 年独立为自治共和国。新加坡华人占 78%，马来人占 15%，印度人占 5%。新加坡宪法规定马来语、华语、泰米尔语、英语 4 种语言为官方语言。马来语是马来人的民族共同语，为尊重历史，新加坡规定马来语为国语，国歌要用马来语唱，以马来西亚和印度尼西亚两国共同的标准马来语作为规范。华语是新加坡华人的民族共同语，以中国大陆的汉语普通话为规范，以北京的语音为标准，同时采用中国大陆的简化汉字和汉语拼音方案。泰米尔语是新加坡印度人的民族共同语，泰米尔语的标准遵循斯里兰卡和印度的原有规范。英语是新加坡行政、教育和贸易用语，是新加坡各民族交际的共同语，以英国 BBC 广播公司的发音为标准。新加坡实行的这种以英语为主流，马来语、华语、泰米尔语并存的多语制，对社会政治经济发展和民族团结等问题都起到了非常重要的作用。

印度独立于 1947 年，1950 年成立印度共和国。印度共和国是一个多民族、多语言、多文字的国家，建国之初遇到了严重的语文问题。印度政府十分注重语言文字的地位规划。印度宪法规定，印度有 14 种法定语言，其中天城体字母书写的印地语是惟一的联邦官方语言，这一规定在 1950 年印

度共和国成立15年之后正式施行。在这15年中,同时也使用英语。而且英国给印度的大量教育拨款用于英语教育,在很长一段时间内,英语一直是印度的行政语言和教育语言。1963年印度国会决定延长英语使用的期限,一直到不需要英语时为止。这样,尽管宪法中规定的14种法定语言中并没有英语,但实际上英语是印度的官方语言。除天城体书写的印地语之外,梵文和乌尔都语也是不限定流通范围的法定语言。梵文是印度教的古典文学语言,代表印度的传统文化。乌尔都语用波斯阿拉伯字母书写,词语有许多来自波斯语,是伊斯兰教的文字,尽管乌尔都语已成为巴基斯坦的国语,但印度仍有3 200万人使用乌尔都语,而巴基斯坦仅有600万人使用。除此之外,印度还有11种限定地区的法定语言,它们是印度各邦的官方语言:古吉拉特语、马拉地语、奥里雅语、孟加拉语、阿萨姆语、旁遮普语、泰米尔语、克什米尔语、泰卢固语、马拉亚兰语和坎纳达语。印度还有11个邦和9个直辖区没有自己本邦的法定官方语言,它们都是用印地语。印度在使用英语的同时,还花很大的力气来扶持印地语的发展和完善,以便将来能够取代英语,成为印度的联邦官方语言和教育语言。像印度这样多民族、多语言、多文字的国家,语文生活长期处于"语言异声、文字异形"的状态,对印度政府来说怎样确定各种语言和文字的地位,怎样制定好语言文字的地位规划,是一个十分重要的问题。

文字地位规划是语言规划的另一重要内容。

从古罗马时代开始,欧洲就掀起了拉丁化的浪潮,拉丁字母成为罗曼语族的文字,意大利语、法语、西班牙语、葡萄牙语、罗马尼亚语都采用了拉丁字母。拉丁字母逐渐代替了原来的鲁纳字母(rune),成为日耳曼语族的文字,英文在5～6世纪使用的也是鲁纳字母,后来也改用拉丁字母,这样德语、荷兰语、卢森堡语、丹麦语、挪威语、瑞典语、冰岛语都采用了拉丁字母。拉丁字母又向东欧扩展,波兰语、捷克语、斯洛伐克语等也都放弃了斯拉夫字母采用了拉丁字母。此外,阿尔巴尼亚语、立陶宛语、拉脱维亚语、爱尔兰语等也都采用了拉丁字母,地中海中的马耳他也采用了拉丁字母。随着新大陆的发现,拉丁字母也传到了美洲,整个美洲都采用了拉丁字母,南美洲和中美洲完全拉丁化了,因此人们又把南美洲和中美洲称作拉丁美洲。

第二次世界大战后,拉丁化浪潮席卷大洋洲,使整个大洋洲都拉丁化了,而且一些新独立的国家也都采用英语作为官方语言,本民族的语言也都逐渐采用了拉丁字母。非洲以撒哈拉沙漠为界,沙漠以北的阿拉伯国家采

用阿拉伯字母，以南除了埃塞俄比亚采用传统的民族字母外，都采用了拉丁字母。亚洲从地理位置上看，西部使用阿拉伯字母，南部使用印度字母，东部使用汉字。第一次世界大战后，作为亚洲国家的土耳其，1924 年就宣布放弃原来使用的阿拉伯字母，采用拉丁字母。第二次世界大战后，印度尼西亚、马来西亚、菲律宾冲破了印度字母和阿拉伯字母的传统，采用拉丁字母。越南本来使用的是汉字，独立后也采用了拉丁字母。我国 1958 年在第一届全国人民代表大会第五次会议上一致通过了拉丁化的《汉语拼音方案》，通过先学习汉语拼音字母，再帮助学习汉字，在推广普通话和对外汉语教学等方面起到了较好的效果。

现在全世界已有 120 多个国家采用拉丁字母为正式文字，不用拉丁字母为正式文字的国家也有拉丁字母的拼写法，作为辅助文字或拼音符号。

**2. 语言文字的本体规划**

语言文字的本体规划，就是对语言文字本身的改造和完善，也就是人们通常所说的语言文字的规范化和标准化。在语言文字的本体规划中，主要是要保护好语言文字本身的纯洁和健康，要对语言文字本身进行精心的加工和培育。语言的本体规划主要包括以下三方面内容：

(1) 推广全民族共同语并使之规范化

推广全民共同语是语言本体规范化最重要的方面。日本是民族相对单一的国家，方言之间的差异较小。明治维新后只用了 20 年的时间就普及了以东京话为标准的国语。日本主要采取了以下两条主要手段：一是重视教育，把国语定为教学语言；二是通过在汉字旁加注假名的办法，达到读音标准化，从而促进国语的普及。正因为日本能够迅速推广民族共同语，才使得日本得以迅速发展。因为日本人已经较早地意识到“共同语是教育的前提，而教育则是建国的前提”。

我国从清朝末年开始开展国语运动，1909 年清政府设立“国语编查委员会”，1911 年通过《统一国语办法案》。五四运动后，国语开始推行，修订注音字母，制定国语罗马字拼音法式，调整国音标准，推动学校的国语教育，培训国语师资，出版国语书刊。但究竟什么是国语还没有一个明确的概念。1949 年新中国成立后，对国语作了明确的界定，改称为普通话，即汉民族共同语。普通话的概念是以北京语音为标准音，以北方话为基础方言，以典范的现代白话文著作为语法规范的汉民族共同语。1956 年，国务院发了《关于推广普通话的指示》，成立了普通话推广工作委员会。把推广普通话的工

作作为一项重要任务写进了周恩来总理的《政府工作报告》中，在全国范围内从中央到地方，从语言研究机构到文化、教育、出版、宣传机构一起行动，充分发动群众，掀起了一场遍及全国的学习普通话的热潮。1957年11月发表了《汉语拼音方案》，1958年2月全国人民代表大会批准了这一方案，这对普通话教学有很大推动。1958年到1964年，推广普通话工作委员会和教育部等单位先后联合举办了4次全国普通话成绩观摩会。“文革”期间推广普通话工作基本上没有进展，“文革”结束后，普通话推广工作在全国逐渐开展起来。1978年5月，中国文字改革委员会在江苏省吴县召开南方方言区推广普通话工作座谈会，8月教育部发出了《关于加强学校普通话和汉语拼音教学的通知》。为培训推广普通话的工作人员，1979年2月和10月，教育部、中国文字改革委员会和语言研究所举办了两期普通话研究班。1978年底到1979年夏天，湖北、河南等11个省先后举办了普通话教学成绩观摩会。1979年8月，教育部、中国文字改革委员会、共青团中央联合在北京召开了第五次全国普通话教学成绩观摩会，在全国，掀起了又一次学习普通话的热潮。1982年“国家推广全国通用的普通话”条文写进了宪法，推广普通话工作有了法律依据。2000年10月31日中华人民共和国第九届全国人民代表大会常务委员会第十八次会议通过了《中华人民共和国国家通用语言文字法》。该项法律共分总则、国家通用语言文字的使用、管理和监督、附则四章二十八条，为语言文字的规范和使用提供了法律保障。

全民共同语的规范化是语言本体规划的重要组成部分。方言也是需要规范化的，但其重要性是不同的。这里我们主要讨论共同语的规范化。全民共同语的规范化既是一项群众性很强的社会实践活动，也是一项学术性很强的理论研究。

语言规范和语言规范化是两个不同的概念。语言规范是指语言系统在相对稳定的状态中所提供的能够正确表达而又为大多数人所接受的语言形式。而语言规范化则是指人们为谋求和保证信息传递效果的一种主动行为，是人们对语言使用情况和发展情况进行的人为干预，以达到语言交际能按照人们预定的愿望发挥最大的作用。语言规范化最重要的问题在于语言的不断发展和语言使用标准的稳妥之间的矛盾。

下面我们对我国的语言规范问题作一个简单的介绍。

我国自古就对正确使用语言十分重视。首先，他们建立起一种接近于民族共同语的“雅言”(《论语·述而》)。其次，他们重视书面语言在读音、释

义、字形方面的指导工作，朝廷专门设置官员主管。最后，他们还编纂了正义、正字、正音的《尔雅》、《说文解字》、《切韵》等一系列大量的用以匡谬正俗的工具书。我国古代的语言规范可以以"雅正"来概括。所谓雅正就是合乎雅音而又能在字音、字义、字体各方面都没有错误。在这种雅正规范观念的指导下产生了先秦诸子及《史记》、《汉书》等彪炳千古的经典著作。这是雅正规范观念的巨大贡献。

但是，后世的文人把这种雅正看成是一成不变的标准，脱离了实际的活的语言，僵化的结果走向了它的反面。直到清朝末年，追求"言文一致"的白话文运动之后，这一格局才出现了转变。但当时对白话文的规范是什么还是不清楚的。真正把现代汉语书面语的规范提出来，并经过中国有史以来规模最大的语言工作者的集体讨论，对有关规范问题作出明确答案的这一历史任务，是在新中国成立后的 20 世纪 50 年代才完成的。建国后，政府高度重视语言运用问题，而且从白话文运动到当时，我国已经产生了大批典范的白话文著作和一些著名的现代语言学巨匠，研讨现代汉语规范问题的时机已经成熟。1951 年 6 月 6 日《人民日报》发表了社论《正确地使用祖国的语言，为语言的纯洁和健康而奋斗!》，号召人们向中国历史上的语言巨匠和运用现代汉语的大师们学习，正确地使用语言，同语言中的错误现象作斗争。与此同时，《人民日报》还发表了著名语言学家吕叔湘和朱德熙合写的《语法修辞讲话》。《语法修辞讲话》在《人民日报》连载达数月之久，在国内引起了极大的反响。各种语文刊物积极配合，掀起了全国范围的学习普通话的热潮。经过数年的学习、实践、准备，语言规范问题逐渐明确起来。1955 年 10 月中国科学院哲学社会科学部召开了全国性的"现代汉语规范问题学术会议"。这是中国有史以来规模最大的有关语言问题的学术会议，集中了全国各地的语言学家以及与语言问题有关的各界代表，进行了热烈的讨论。会议中心明确、成果突出，对汉语的规范和发展来说，其作用是划时代的。罗常培、吕叔湘的学术报告《现代汉语规范问题》，总结了历史经验，第一次把语言规范问题提到语言学理论的高度加以研究；第一次给汉民族共同语——普通话作出明确的界定，并对普通话形成的历史渊源作出了科学的论证；第一次阐述了语言规范化对发展科学、繁荣文学、普及教育等方面的重大意义；第一次系统地说明了关于语言规范化的理论原则问题。会议结束之后，汉语规范化工作立即在全国普遍展开，短时期内就取得了前所未有的巨大成果。主要体现在全国大力推广普通话、大规模方言调查、制定

汉语拼音方案、整理并简化汉字、审定异读词、在全国各大中学校里加强汉语规范化的教学等等。

当然我们的语言规范化问题也存在着一些不足之处。主要表现为以下几点：

1）语言规范化的工作仅局限在匡谬正俗方面，忽视了前瞻性的引导工作。

2）没有从理论上提出对变化中的语言进行抉择的原则标准，即对突破现有规范的语言现象如何确定它是否合乎语言发展的原则标准。

3）对语言是不断变化发展的认识不足，过分强调语言稳定性的一面，对一些社会需要的、新生的语言形式也加以排斥。

4）虽然在认识上强调“规范化的主要对象是书面语”，但是在实际工作中未能很好地加以体现，主要表现为：不能按照法律的规定行事，政府行为少，投入的力量远远不足等等。

目前我国的语言规范化工作亟需引导和加强，许多报纸、杂志、电视广播等媒体的语言使用质量良莠不齐。造成这一局面的主要原因是有关领导部门对语言运用的质量重视不够，不依法办事，解决问题的力度不够。目前我们亟需做好以下几方面的工作：

1）严格按照《中华人民共和国国家通用语言文字法》的规定对语言文字的使用进行规范。

2）加强正确使用语言文字的宣传工作和引导工作。

3）加强对新闻、出版、教育等机构语言运用的管理工作。

4）成立专门的语言规范化的组织，加强学术研究，提高相关人员的语文素养。

5）加强普通话推广工作，提倡运用规范语言，国家公务人员等要有明确的语文能力要求。

(2) 文字规范和标准的制定

日本在文字生活中较早地注意到了汉字的规范化和标准化。为了解决汉字读音的混乱局面，日本创制了“假名”。1946 年公布了“当用汉字表”共 1 850 个字，1981 年公布“常用汉字表”共 1 945 个汉字，其中简化汉字 225 个。日语的假名有平假名和片假名两种，在使用过程中假名的字形逐渐规范，两种假名也有了分工：平假名趋于通用化，片假名趋于专用化，科技术语、外国的人名地名等都用片假名写。由于假名应用范围的扩大，日本开始

主张不用汉字,全用假名,称作“假名运动”。日本的罗马字运用开始于明治维新前后。1866 年提出《汉字废止议》,1869 年提出《修改国语论》,1872 年提出《文字改换议》,1874 年发表《用洋字写国语论》,1885 年成立“罗马字会”,1905 年成立“罗马字推行会”,1909 年成立“日本罗马字社”。二战以后,1947 年文部省发表《罗马字教育指针》,规定学校罗马字教育实施要项。1949 年成立“日本罗马字教育研究会”,罗马字逐渐成为与传统文字并用的一种辅助文字。日本罗马字有“黑奔式”和“训令式”两种,其中训令式是根据日本语特点设计出来的,并且已经提交国际标准化组织文献标准化技术委员会。在日本文字的计算机处理中,日本采用“假名-汉字互换”的方法,制成了日文文字处理机,成功地解决了日文输入计算机的问题。

我国的文字规范和标准的制定可参看本节“六”。

(3) 科学技术术语的标准化

这部分内容请参看本节“七”。

## 三、文字改革

文字改革是指文字体系和文字制度的局部或整体变革。文字改革是一种非常重要的语言规划任务,它既包括语言文字的地位规划也包括语言文字的本体规划。文字改革的形式是多种多样的,主要有以下两种情况:一是创制文字,就是一个国家或一个民族在没有自己文字的情况下,创制属于自己的新文字,或者是在创制自己的文字前,一直使用其他民族或国家的文字,由于社会的发展和需要创制属于自己的文字;一是对业已存在的文字体系进行局部的调整和改革,包括对字母和字符的增添等。

文字改革一般而言有以下几种原因:首先是经济的发展要求提高语言效率;其次是人口的迁徙,需要选择和推广通用的文字;三是新独立的国家需要创建新的属于自己的文字;四是新术语和语文科技的出现与发展,需要对原有的文字体系或古老文字进行改革。

中国的文字改革主要指汉字的改革,当然也包括一部分少数民族的文字改革。我国的文字改革历史悠久。比较近的有 1892 年卢戆章的《切音新字》、北洋政府教育部制定的一套注音字母、1928 年南京政府公布的国语罗马字拼音法式、1931 年由留苏的共产党员和苏联语言学者设计的一套中文拉丁化字母(称为拉丁化新文字)等等。建国后,我国成立了中国文字改革委员会。1956 年政府公布了“汉字简化方案”。同年又提出了我国文字改革的 3 项任

务:简化汉字、推广普通话、制定和推行汉语拼音方案。在我国文字改革与语言规划是密不可分的,我们所说的语言规划中就包括文字规划。

## 四、汉语拼音运动与汉语拼音方案

### 1. 汉语拼音运动

明朝末年西方传教士到中国传教,为了学习汉字,他们开始使用拉丁字母拼写汉语。1605 年意大利传教士利玛窦出版了《西字奇迹》,其中有 4 篇文章加了拉丁字母的注音,是最早的用拉丁字母给汉字注音的出版物。随后法国传教士金尼阁出版了一本用拉丁字母给汉字注音的字汇——《西儒耳目资》。1867 年英国的威妥玛在北京出版了北京语音官话课本《语言自迩集》,他设计了一套拼写法,用拉丁字母拼写中国人名、地名和事物等的名称,被称作"威妥玛式"。还有一些方言教会罗马字在南方口岸传播。这些用拉丁字母拼写汉字的方法为汉语拼音运动提供了经验和基础。

中国自发的拼音运动是从清朝末年的切音字运动开始的。1892 年,卢戆章出版了《一目了然初阶》,公布了他创制的"中国切音新音",用拉丁字母及其变体来拼写厦门音,增加声母后可以拼写泉州音和潮州音。从此,开始了长达 20 年的切音字运动。在此期间,新的切音字方案层出不穷,其中推行较广的有王照的官话字母和劳乃宣的合声简字。1913 年读音统一会议定了"注音字母",并通过了《国音推行方法七条》。1918 年北洋政府教育部正式公布"注音字母"。1920 年全国各地陆续开办"国语传习所"、"暑期国语讲习所",推广注音字母,全国小学的文言文课一律改为白话文课,小学教科书都在汉字的生字上用注音字母注音。注音字母在我国用了近 40 年的时间,对我国统一汉字读音、推广国语、普及拼音知识起到了非常大的作用。1923 年教育部召开国语统一筹备会,决定组织"国语罗马字拼音研究委员会"。1925 年在北京的部分委员和一些语言学者自动组织"数人会",提出了《国语罗马字拼音法式》。1928 年蔡元培正式公布《国语罗马字拼音法式》,这就是国语罗马字运动。但国语罗马字没有走出知识阶层的圈子,没有在社会上普遍推行。

中国的拉丁化新文字是在 20 世纪初的 20 年代末和 30 年代初在苏联创制的,其目的是扫除在苏联远东的 10 万华工文盲,试图在时机成熟的条件下,用拉丁文字代替汉字。中国的拉丁化新文字运动是瞿秋白倡议设计的,1929 年和 1930 年分别发表了《中国拉丁字母草案》、《中国拉丁化字

母》。1931 年 9 月在海参崴召开的中国新文字第一次代表大会上，通过了书面方案《中国汉字拉丁化的原则和规则》，标志着拉丁化新文字运动正式开始。1935 年，蔡元培、鲁迅、郭沫若等众多文化名人积极拥护拉丁化新文字，发表了《我们对推行新文字的意见》。直到 1958 年《汉语拼音方案》公布为止，拉丁化新文字运动对我国的文字改革、对制定和推广《汉语拼音方案》都起了极其重要的作用。

2. 汉语拼音方案

1949 年中国文字改革协会成立，其中设有专门的"拼音方案研究委员会"，主要讨论采用什么字母的问题。1955 年"拼音方案研究委员会"最终确定拼音方案采用拉丁字母。1956 年 2 月中国文字改革委员会发表《汉语拼音方案(草案)》公开征求意见。同年 10 月国务院设立汉语拼音方案审订委员进行审议，1957 年 10 月提出修正案，11 月 1 日由国务院全体会议第 60 次会议作为《汉语拼音方案(草案)》通过，提请全国人民代表大会讨论。1958 年 2 月 11 日，《汉语拼音方案》由第一届全国人民代表大会第五次会议正式通过。《汉语拼音方案》通过后，在全国范围内得到了迅速的推广和应用。1958 年秋季开始，全国小学的语文课本一律采用汉语拼音给汉字注音，扫盲课本也使用汉语拼音注音，对外国人学习汉语也起到了积极的作用。汉语拼音字母成为字典和词典注音、排序的标准，一些书刊的索引也采用了汉语拼音字母。《汉语拼音方案》的推广和运用还对少数民族的文字改革和创制、盲文和手语的制定提供了依据，在视觉通讯、无线电报、计算机的文字输入和输出、产品型号标记等各行各业均得到了广泛应用。

1978 年 9 月国务院批准《汉语拼音方案》作为我国人名、地名罗马字母拼写法的统一规范。1982 年 8 月国际标准化组织通过决议，采用汉语拼音作为在文献工作中使用罗马字母拼写有关中国词语的国际标准。《汉语拼音方案》已经由中国标准发展成为国际标准，标准号是 ISO7098，但我们也应该清楚地认识到，汉语拼音是汉字拼写的辅助工具，不是拼音文字。

## 五、普通话的推广

"普通话"一词最初是由清朝末年切音字运动的先驱者之一的朱文熊首先提出的。此后，吴敬恒、胡适、瞿秋白、鲁迅、陈望道等都提出了相似的观点。但客观地讲，这一时期内人们心目中的"普通话"，还不是由北京语音为标准音的"国语"。

20世纪50年代初，大陆学者为确定民族共同语，展开了一次关于共同语标准的讨论。会议上有两种主张：一是以北京话为标准，一是以普通话为标准。这两种意见最终都没有成功。普通话获得明确的内涵和共同语地位是在1955年"现代汉语规范问题学术会议"上。罗常培和吕叔湘联合作的《现代汉语规范问题》的报告第一次全面系统地论述了与汉语规范化有关的一系列问题，其中就包括民族共同语的形成条件、形成过程、规范标准和代表性称谓的论述。经过与会代表的认真讨论，认清了汉语规范化的重要意义，明确了汉语规范化的标准是"以北京语音为标准音，以北方话为基础方言，以典范的现代白话文著作为语法规范的普通话"。从此，普通话被赋予了新的涵义，获得了汉民族共同语的标准语地位。

1. 五六十年代的"推广普通话"工作

20世纪五六十年代，国家为推广普通话工作采取了一系列措施，取得了骄人的成绩。1956年2月国务院成立了中央推广普通话工作委员会，并发出《关于推广普通话的指示》，作出了一系列针对各行各业的详细的运用普通话规定。各省市相继成立了推广普通话工作委员会。国务院指示各省、市、自治区在教育厅内设立普通话推广处。中央还制定了"大力提倡、重点推行、逐步普及"的推广普通话的工作方针。从1956年起，全国中小学的语文课本开始教学普通话。1957年11月《汉语拼音方案》发表，1958年2月得到全国人民代表大会的批准，这对普通话的推广起到了积极的促进作用。1958年，毛泽东主席指示"一切干部要学普通话"；周恩来总理在政协全国委员会所作的报告《当前文字改革的任务》中指出：在我国汉族人民中努力推广普通话"是一项重要的政治任务"。

由于各地方言和普通话之间的差别主要表现为语音的不同。做好普通话语音师资培训和各地方言的调查工作，就成了推广普通话工作中一项十分重要的任务。1955年7月，教育部发出了《关于举办小学语文教师标准语语音训练班的通知》。为帮助全国中小学生和方言区的人民群众学习普通话，从1956年开始，中央人民广播电台多次举办了普通话语音教学广播讲座、中小学和师范学校语文朗读广播讲座。1956年至1961年间，中国文字改革委员会先后举办了9期普通话语音研究班，培训普通话师资，各省市也进行了培训工作。为配合普通话的推广工作，高等教育部和教育部于1956年3月发出《关于汉语方言普查工作的指示》，调动各地综合性大学和高等师范院校的中文系，以县、市为方言点着重调查各地方言的语音差异。

在两年多的时间里共普查了 1 849 个方言点，编写出调查报告近 1 200 份。并在此基础上，根据当地方言同普通话之间的语音对应关系，编写出了 300 多种供当地人学习的普通话学习手册，对普通话的普及发挥了重要作用。1956 年成立"普通话审音委员会"，对 1 800 多条异读词和 190 多个地名的读音进行了审议，1963 年出版了《普通话异读词三次审音总表初稿》，为语音的规范化提供了依据。总之，20 世纪五六十年代的"推广普通话"工作，在全国范围内掀起了一个遍及城乡的推广、学习普通话的群众运动，形成了以学习普通话、说普通话为荣的风尚。

2. 新时期的"推广普通话"工作

"文革"期间的"推广普通话"工作受到了严重的干扰和破坏，处于停顿状态。尽管"文革"后期广大语文工作者和教育工作者也作了一些努力，但收效不大。1977 年国务院发出了有关推广普通话工作的文件，1978 年教育部发出了《关于加强学校普通话和汉语拼音教学的通知》。中央人民广播电台举办了汉语拼音广播讲座，文改会又恢复了普通话研究班的培训工作。1979 年 8 月在北京召开了第五次全国普通话教学成绩观摩会。1982 年文改会恢复了普通话审音委员会，1985 年 12 月通过了《普通话异读词审音表》。1982 年 11 月第五届全国人民代表大会第五次会议通过的《中华人民共和国宪法》中，写进了"国家推广全国通用的普通话"，使普通话的推广工作有了法律的依据。1983 年教育部专门召开了中等师范学校推广普通话工作会议。

1984 年是推广普通话工作的一个新的转变期。1984 年 10 月，中国文字改革委员会在北京怀柔召开文字改革工作座谈会。会议对 50 年代提出的语言文字工作和文字改革的任务作了调整，把推广普通话列在了首位，说明当时语文学界对推广普通话的重要性已经有了新的认识。1985 年 12 月，国务院通知：为加强新时期的语言文字工作，将原中国文字改革委员会更名为国家语言文字工作委员会。1986 年 1 月国家教育委员会和国家语言文字工作委员会在北京召开了"全国语言文字工作会议"，这是 1955 年 12 月"现代汉语规范问题学术会议"召开 30 年后，标志中国语言文字工作新开端的又一里程碑。关于新时期的推广普通话工作，会议指出"重点应当放在大力推行和积极普及方面"。会议提出在"本世纪"(即 20 世纪)内，推广普通话的目标是：第一，各级各类学校采用普通话教学，使普通话成为教学语言；第二，各级各类机关进行工作时一般使用普通话，使普通话成为工作语言；第三，广播、电视、电影、话剧使用普通话，使普通话成为宣传语言；

第四,不同方言区的人在公共场合的交往基本上使用普通话,使普通话成为交际语言。关于加强语言文字工作的措施有两点值得指出:一是提出了要健全机构,除国家语言文字工作委员会外,省、市、自治区也要设立相应的机构;二是加强有关的科学研究,取得具有较高水平和实用价值的研究成果。全国语言文字工作会议不仅是新时期以来我国语言文字工作的新开端,也是新时期推广普通话工作进入新阶段的标志,从此推广普通话工作成为新时期语言文字工作的首要任务。1986 年中央六部委联合发出了《关于加强开放、旅游城市推广普通话工作的通知》,1987 年和 1988 年,国家语委先后在大连和青岛召开了沿海开放城市和经济特区推广普通话工作座谈会,1990 年国家语委在上海召开了全国城市社会推广普通话工作经验交流会。

20 世纪 90 年代初,国家语委根据十一届三中全会精神和我国社会主义现代化建设的实际需要,制定了《国家语言文字工作十年规划和"八五"计划纲要》,《纲要》提出"八五"期间首要的工作任务就是"推广普通话",而且指出:"必须大力推行,积极普及,逐步提高,在深度和广度上要上一个新台阶。"1992 年,国家语委根据《纲要》的精神向国务院呈报了《关于当前语言文字工作的请示》,同年 11 月,国务院批转下达了这一《请示》,并要求各级政府"必须予以高度重视"和"贯彻执行","坚持不懈地抓好推广普通话、推进文字规范化、推行汉语拼音等工作,使语言文字工作更好地为社会主义现代化服务"。1992 年至 1993 年,国家教委和国家语委进行了全国中师和高师普及普通话的检查验收工作。1994 年国家教委和语委制定颁发了《师范院校普通话工作评估指导标准》和《城镇中小学普及普通话工作评估指导标准》,标志着学校普及普通话工作开始进入制度化、规范化、科学化的阶段。同年,国家语委、国家教委和广播电影电视部联合发出《关于开展普通话测试工作的决定》,并颁布了《普通话水平测试实施办法(试行)》,规定了现阶段的主要测试对象和应达到的级别标准,对达到相应级别的人员经主管部门复议后发给等级证书,对必须使用相应级别的普通话的工作逐渐实行持普通话等级证书上岗的制度。2000 年 10 月 31 日中华人民共和国第九届全国人民代表大会常务委员会第十八次会议通过了《中华人民共和国国家通用语言文字法》,并于 2001 年 1 月 1 日正式实施,为推普工作、语言文字的规范使用提供了法律保障。

**3. 双语制及双语区和方言区的普通话推广工作**

双语制是指在一个言语社团里,使用两种或两种以上语言作为共同

语，这种现象又叫做双重语言制。双语现象一般是在语言融合过程中才出现的一种语言现象。在许多国家中，当两种语言在相互争夺主导地位的过程中，一种语言一时还不能战胜另一种语言时，就会造成双重语言现象的出现。目前加拿大（英语和法语）、菲律宾（英语和菲律宾语）、瑞士（德语、法语、意大利语和罗曼希语）等等很多国家都实行双语制。双语制作为一种语言融合过程的必然阶段和产物，它的发展结果只有两种：一是两个民族向融合的方向发展，相互间的关系越来越密切，其中一个民族主动放弃自己的语言，完成语言的融合；一是两个民族向分离的方向发展，继续使用各自的语言。双语现象的产生有历史和社会等多种因素。我国的民族自治地区享有法律规定的使用本民族语言的权力，这些地方的推普工作在积极引导的同时，还要有相应的保持该民族自己特有的语言文字的措施。

在一个具有方言分歧的社会里，人们在交往时往往选择其中的一种方言作为相互交际的"通用语"，这种"通用语"是方言区之间的交际工具。这些带有地方色彩的"普通话"，是一种介于地方方言和普通话之间的一种中间状态的语言现象，是一种过渡性质的语言。这种地方普通话吸收了标准普通话的词汇和语法，但语音或标准音和方言音两者混杂，还有一些成分既不是地方普通话中的，也不是标准普通话中的。地方普通话是标准普通化存在的主要形式，也是丰富和发展标准普通话的一个重要的源泉，因此在实行语言规范化和推广普通话的过程中，无论是从丰富和发展标准普通话来看，还是从普通话的规范和标准来看，地方普通话都是一种非常值得研究的语言现象。

民族共同语也是在一种方言的基础上产生和发展起来的。在一个多民族的国家中各民族之间还需要有一个共同的交际工具，当国家通过法律确定了各民族在政治、经济、文化、教育等方面共同使用的一种语言作为交际工具时，这种语言就成了这一国家的"国语"，简而言之，国语就是一个国家内部各民族之间共同的交际语言，国语可以是一种，也可以是两种或两种以上。如我国是汉语，加拿大是英语和法语，瑞士是德语、法语、意大利语和罗曼希语等等。国语在我国有特定的含义，专指五四前期，"国语运动"要把北京话作为汉民族共同语，着眼于口语，想用国语代替方言，实现语言的统一，这里所说的国语实际上就是普通话。台湾省目前仍把普通话叫做国语。

## 六、汉字的整理、简化和标准化

### 1. 汉字的整理

汉字的整理主要是为汉字定音、定量、定形、定序，使之能够更规范标准，更适于运用。汉字的整理工作由来已久，春秋战国时期的《史籀篇》、《仓颉篇》等就规定了儿童学习汉字的字数；秦统一汉字初步确定了汉字的书写形式(字形)，东汉许慎《说文解字》确立了解释和书写汉字字形的权威；1932年出版的《国音常用字字汇》收正字 9 920 个，别体重文 1 179 个，变音重文 1 120个，并确立了以北京语音为字音的标准，加收了一些通用的简体字形，排序按注音字母排列，在字形、字音、字量、字序等方面建立了初步的规范，是新中国建立前汉字整理的一块里程碑。

建国后，1952 年教育部就公布了《二千常用字表》，1955 年 12 月文化部和文改会联合公布了《第一批异体字整理表》，翌年国务院公布《汉字简化方案》，1964 年编制了《简化字总表》。1956 年中国科学院语言研究所成立普通话审音委员会，先后发表了《普通话异读词审音表初稿》正编、续编和第三编，1985 年发表修订稿《普通话异读词审音表》。1955 年文改会组成标准字形研究组，1960 年拟定《通用汉字字形表(草案)》，1964 年公布《印刷通用汉字字形表》。1961 年，文化部、教育部、文字改革委员会和中国科学院语言研究所联合组成汉字查字法整理工作组，对音序、部首四角号码和笔画笔形等 4 类检字法方案进行讨论研究，并提出这 4 类检字法的方案草案，1983 年 6 月拟定出《统一汉字部首表》(征求意见稿)。1980 年 5 月，中国文字改革委员会提出并通过了关于研究和制定《标准现代汉语用字表》的建议和科研计划草案。1986 年 1 月全国语言文字工作会议把“研究和整理现行汉字”定为当前语言文字工作的一项重要任务。

### 2. 汉字的简化

汉字简化是专门就汉字的形体进行的简化，包括笔画的简化、精简异体字等。汉字的简化是随着社会的发展而不断进行的。甲骨文时期汉字中就有异体字，笔画的简化是产生异体字的一个主要原因。小篆、隶书、楷书、行书等都曾有过正、草的分化，而且草体始终走在正体的前面。1909 年费孝通在《教育杂志》创刊号上发表《普通教育应当采用俗体字》一文，提倡使用简体字。1922 年钱玄同和黎锦熙、杨树达在国语统一筹备委员会上联名提出《减省现行汉字的笔画案》，对简体字运动起到了积极的推动作用。1935

年8月,"国民政府"教育部公布第一批简体字表,是第一份由政府部门正式公布的简体字表。解放后,1950年中央人民政府教育部社会教育司编制《常用简体字登记表》征求意见,1951年拟定出《第一批简体字表》。1954年中国文字改革委员会在《第一批简体字表》的基础上拟出《汉字简化方案(草案)》。1955年《汉字简化方案(草案)》发表并开始试运行。1955年国务院成立汉字简化方案审订委员会。1964年中国文字改革委员会编辑出版了《简化字总表》。我国《第二批汉字简化方案(草案)》没有推行开来。国外的一些使用汉字的国家和华人社区也根据祖国大陆进行了相应的汉字简化工作。汉字是需要简化的,但这并不意味着汉字越简单越好,在一定的时期内还要保持汉字的稳定,这才有利于汉字的运用和研究,为可能的简化提供理论和实践上的准备。

**3. 汉字的标准化**

在计算机尚未广泛运用时期,汉字的简化主要是为了便于认读、书写,为普及文化教育、扫除文盲和排版印刷提供便利。随着我国科技的发展,特别是计算机时代的到来,对中文信息处理提出了许多新的问题,同时也为汉字研究赋予了新的内涵。

汉字的数量是非常庞大的。1988年国家语言文字工作委员会和新闻出版署发布的《现代汉语通用字表》共收字7 000个,可看作是现代汉语通用汉字的定量标准。同年发布的《现代汉语常用字表》收字3 500个,可看作常用汉字的定量标准。从汉字的使用频率看,其中最常用的前1 000个字对文章的覆盖率为90%,前2 400个字的覆盖率就会达到99%。汉字虽然从字符上看相当庞大,但是如果从造字方式上看,汉字有自己特殊的性质。以前汉字主要采用"六书"的方式造字,但目前现代汉字的字形构造有的已经看不出其理据来了。现代汉字的形体可以分成3个层次:字、部件和笔画。一个字由几个部件构成,一个部件又是由几个笔画构成的。一个笔画最繁复的汉字需要几十个笔画符号,但这些笔画有相当多是重复的,实际使用的笔画数是不多的,而这些笔画可以只需要几个,最多十几个。这又体现出了汉字字形构件的简洁性。

在信息时代,怎样把汉字输入输出计算机已成为中文信息处理的关键问题之一。我国从20世纪60年代就开始对汉字信息处理进行探索和实践,1968年研制成汉字电报译码机,70年代中期明确提出了"汉字信息处理系统"的研究课题。1978年以来,中国开始广泛应用大规模集成电路存储

器和成套的微处理机芯片,为汉字输入计算机提供了物质基础,成功研制了一些新的汉字输入输出设备,并配制成各种应用系统。目前我国的汉字输入方法大致可分为6类:编码输入法、整字输入法、拼音-汉字转换法、印刷体光学输入法、手写输入法和语音输入法。

汉字的输入法可以在众多输入法中选出最佳的汉字编码方案,但同时汉字的输出也是汉字信息处理的难点之一。为了解决这一问题,"汉字信息处理系统"把能够输出高质量汉字的汉字照相排版系统作为重点攻关项目,经过努力终于取得了令世人瞩目的成就。目前我国普遍采用的是计算机激光汉字编排系统,具有中文报刊的国家纷纷从我国订购这一照排系统。这对我国今后计算机信息处理起到了非常关键的作用,使我国在这一领域处于世界领先地位,为迎接信息化社会的到来打下了坚实的基础。

计算机汉字信息处理中常用汉字的标准化和规范化也是极其重要的。我国制定了《信息交换用汉字编码字符集》的国家标准。1981年又公布了国家标准《信息交换用汉字编码字符集——基本集》(GB2312-80)。1984年全国计算机与信息处理标准化技术委员会字符集和编码分技术委员会经过研究,提出了汉字编码字符集标准的繁体字和简体字对应编码的原则,并做出了制定6个信息交换用汉字编码字符集的计划。

目前除中国以外,还在使用汉字的国家有:韩国、日本、新加坡、马来西亚等。马来西亚和新加坡目前都已使用简体字,其他的国家和地区还没有统一标准,这在信息交流的过程中带来了极大的不便,有必要建立一套统一的汉字字符集。我国大陆除了制定了《信息交换用汉字编码字符集》外,还制定了《信息处理文本通信用编码字符集》(GB 8586);台湾制定的字符集标准主要有两个,一个是《中文资讯交换码 CCC Ⅱ》,一个是《通用汉字标准交换码 TCA-CNS 11643》。韩国1987年制定了韩国标准编码字符集 KS C 5601-1987,并于1991年又制定了辅助集 KS C 5657-1991。日本制定的字符集标准主要有 JIS X 0208-1983 和 JIS X 0212-1990。国际标准化组织于1984年开始研究制定《信息技术通用多八位编码字符集(UCS)》国际标准。1993年5月该标准的第一部分"体系结构与基本多文种平面"正式发布。它对信息技术通用多八位编码字符集(UCS)的总体体系结构和基本多文种平面作了规定。规定在基本多文种平面内,A区用于字母文字、音节文字和各种符号的编码,I区用于中日韩统一的表意文字编码,O区留作未来标准化用,R区作为限制使用区,用于专用字符、变形显现和兼容字符的编码。I

区中的表意文字采用中日韩汉字统一编码方式，以现有的各种标准字符集为源字符集，将其中的汉字按统一的认同原则进行认同和甄别后，生成涵盖源字符集，并按东亚著名的四大字典《康熙字典》、《大汉和字典》、《汉语大字典》和《大字源》的页码/字位综合排序的 DJK 汉字统一字汇和字序，构成信息技术通用多八位编码字符集中的表意文字部分，共包括 20 902 个汉字字符。综合各国的情况看，中日韩汉字统一编码可以满足中国(含台湾)、韩国和日本汉字信息处理的需要，在国际信息处理中是一件非常有意义的大事。从此结束了长期以来中日韩汉字统一编码混乱的局面，对促进汉字信息处理的进一步发展，为推进整个信息产业的进一步发展奠定了良好的发展基础。

### 七、术语的标准化

术语是人类社会科学技术发展的必然产物。术语是语言的一部分，对术语进行必要的规范化和标准化也是语言规划中本体规划的一个重要组成部分。人类在早期的哲学、自然科学和人文科学等领域就开始注意到术语的使用问题。18 世纪以来，随着科学技术的迅速发展，学术交流越来越频繁，新的学术概念层出不穷，代表这些学术概念的术语也越来越多，术语的命名和统一就成了科学家们探讨的重要问题。但当时关注这一问题的大多是自然科学家，语言学家很少。19 世纪后，一些国家建立起了各个学科的科学技术协会，它们都不同程度地开展了有关学科的术语研究工作。1867 年在国际植物学家大会上提出了植物学术语命名的统一规则，1889 年国际动物学家大会上提出了动物学术语命名的统一规则，1892 年国际化学家大会上也提出了化学术语命名的统一规则。这时，语言学家才开始关注术语的研究。20 世纪初，出现了协调术语的国际性组织。1936 年国际标准化协会建立了第三十七技术委员会，1951 年国际标准化组织(ISO)继承了术语学委员会的工作，建立了术语学原则和协调委员会。这些组织的成立都有力地推动了术语标准化的工作。

奥地利科学家维斯特 1931 年写成了第一篇关于术语学的论文《在工程技术中(特别是电工学中)的国际语言规范》，文中提出了现代术语学的基本原则和方法，阐述了术语系统化的指导思想，为现代术语学奠定了理论基础，随后他又发表了一系列有关术语的论文。维斯特因此被誉为现代术语学的奠基人。现代术语学可以分为 4 个学派：德国-奥地利学派、俄罗斯学派、捷克斯洛伐克学派和加拿大-魁北克学派。

1963年卢森堡建立了世界上第一个术语数据库,并用计算机进行管理,为术语数据的存储、检索提供了方便。目前,德国、卢森堡、法国、俄罗斯等国家都建有术语数据库,这些数据库的建立标志着现代术语学的发展进入了一个新阶段。1971年维也纳建立了国际术语信息中心,该中心与国际标准化组织合作,开展术语学人才的培养、国际学术交流,召开国际学术会议,并在国际范围内协调术语学研究和术语工作,它还致力于国际术语网络的建立工作。

我国古代的科学技术是十分发达的,与此相应的就是术语的大量出现。对术语的关注自然很早就引起了一些人的关注。比较有代表性的论述是荀子的《正名篇》,书中提出了很多语言学理论方面的问题,而且许多论点都与术语有关,其中很多对今天的术语学研究仍然具有一定的启发意义。以后历代也都有许多涉及各种术语的典籍和著作,但总体来看,术语的使用还是十分混乱的,使用的范围也比较有限,不利于相互交流、学习和借鉴。到了明末清初,人们开始意识到术语统一的重要性,开始按照西文第一音节创造中文的化学元素术语,开始采用音译和意译的方法翻译西方的术语。到了清朝末年,要求统一术语译名的呼声越来越高。1909年,成立了以严复为总编纂的科学名词编订馆,这是我国第一个审定科技术语的统一机构。辛亥革命后,江苏教育会的理化教授研究会首先审定了物理学和化学术语,中华医学会组织了医学名词审查会,1915年相继审定了化学、物理学、数学、动物学、植物学、医学的术语。1919年成立了科学名词审定委员会。1932年成立了国立编译馆,专门负责管理全国科技术语的审定工作。新中国成立的第二年,成立了学术名词统一工作委员会,下设自然科学组、社会科学组、医药卫生组、时事文学组、艺术组。1956年学术名词的规范化工作由中国科学院负责,在中国科学院编译出版委员会下设了名词室,负责统一和审定全国自然科学术语的工作。"文革"后,1985年4月全国自然科学名词审定委员会成立,其主要任务是:确定工作方针,拟定全国自然科学名词统一工作计划实施方案和步骤,负责审定自然科学各学科名词术语的统一名称,并予以公布。全国自然科学名词审定委员会按数学、物理学、化学、天文学、地球科学、生物科学、技术科学、农业科学、医学以及交叉科学等专业领域,分别成立专门的名词审查分委员会,负责制定本学科术语审定与统一的近期和长远计划,并开展各种审定活动,取得了巨大的成绩。近年来,我国社会科学也开始重视名词术语的审定工作,如2001年成立了语言学名词审定

委员会，负责语言学名词的审定工作。

标准化不仅仅只局限于产品的标准化，术语也是标准化的一项重要内容，而且从规范和发展的角度看，术语的标准化应该先于产品的标准化。

为推动我国术语的标准化工作，国家技术监督局建立了全国术语标准化技术委员会，下设3个分委员会，分别负责术语学的理论和应用、辞书编纂、计算机辅助术语工作，目前已经完成了多项国家标准。1993年我国还成立了中国术语工作网，同时我国也开始了术语数据库的研制工作，并于1985年建立了一个英汉对照的术语数据库。我国的术语工作还有许多亟需解决的问题，如只注重应用，缺乏理论建设；大陆（包括港澳）和台湾间的术语差异过大等等，都给科技和经济的发展带来了负面影响。我们应当抓紧机遇努力缩小术语研究与发达国家间的差距。

## 第二节 语言调查

语言是一种非常复杂的社会心理现象。要想全面地了解语言的面貌，就必须对语言进行详细、系统的调查，特别是要对语言内部语音、词汇、语法等的特点及各种内部差异，亲自到各地去口问、耳听、手记，作语言静态和动态的描写工作，这就是语言的田野工作，也是语言调查的基本的工作方法。

### 一、基本程序和方法

#### 1. 科学研究所遵循的原则

对一种语言现象的思考，人们首先形成的是一些初步的看法和假设，然后再试图说明这一现象和相关现象之间的关系。要想证明自己的假设是否符合实际，人们就必须设计一个研究方案，制定出实施的步骤，通过有目的的观察去收集相关资料，通过对资料的分析和归纳、抽象和推断，得出结论。由此可见，科学研究必须依靠科学的研究方法。在语言研究中必须遵循科学的研究原则是：

第一，收集和解释的资料或数据必须是真实的、客观的，不能带有主观臆测的或偏见的成分。

第二，综合和推理必须有理有据，合乎逻辑。

第三,研究方法要公开整个程序,结论要经得起实践的检验。所以对运用的程序、所测的变量和使用的度量方法等必须予以详细说明,对提出的命题、假设、使用的概念或术语必须给出清楚的界定与明确的涵义,以便其他研究人员能够依照同一程序、同一方法核实其数据,检验其成果。

社会语言现象的研究不能仅凭个人的直觉和内省体验,或依据少量调查的数据和情况作出断言。20 世纪 60 年代,拉波夫运用定量分析的方法研究语言变量与社会变量之间的关系,其代表作《纽约市英语的社会分层》,为这一领域奠定了基础。这项研究在方法上有两大突破:一是在确定调查单位方面,拉波夫运用了社会学的研究方法,在已有的普查基础上,采用随机抽样法,抽取调查单位,并用社会变值作为划分调查对象层次的尺度;二是在收集言语材料方面,拉波夫设法诱导和分离出 4 种随情景而异的语体,并用录音机记录素材。这两项创举使他能够对调查对象的社会特征作出较周密的分析,掌握大量的言语资料和幅度较宽的语体,从而有条件对社会方言进行大规模的定量分析。

2. 语言调查的基本步骤和方法

后人对拉波夫的调查方法进行了补充和完善,已经形成了一套比较完善的调查体系。下面我们就介绍一下基本的步骤和方法。

(1) 选择课题

课题的选择可以从以下两方面入手。

1) 从自己的感受中去发掘值得研究的言语现象。人们在研究时,一般总是选择自己熟悉的语言变体作为研究的内容。但值得注意的是,并不是所有的语言差异都具有社会意义。拉波夫认为有 3 种类型的差异应当加以区分:一是不带社会意义的差异,如 cot(kat)和 caught(kɔt)两词中的元音,有的美国人加以区别,有的不区别,都发成(a)音。这类差异只起指示作用,不随语境而变;一是具有社会意义的差异,如 car 和 cart 中的 r 音,在纽约市发不发卷舌音体现说话人的社会身份,人们对此十分敏感。这类差异起到标记作用,而且随语境而变。还有一种差异已经定型,具有一成不变的评估标准,如纽约市鲁克林区的居民把 thirty-third 说成 toity-toid,这种说法被视为粗俗的形式,人们对它嗤之以鼻。就变异研究要达到的理论目标来看,第一种和第三种类型的变异研究的价值不大。只有第二种类型的变异是随着说话人的社会属性及说话时的语境而变化,人们对不同的变异形式持有不同的价值标准,因此可以从中发现某些普遍性的规律。

2) 从前人研究的成果中汲取营养，寻求可研究的课题。20 世纪 50 年代，费舍尔曾就变量作过研究。英语中动词的现在分词，如 fishing，hunting 等词尾的语素 ing 有两种读音：(ŋ)或(n)。前者是标准读音，后者是非标准读音。他对 12 名男孩和 12 名女孩进行了调查，发现在正式场合男孩比女孩发非标准读音的人要多，模范男孩比调皮男孩发标准读音多。此外他还注意到孩子们在轻松的时候发非标准读音的次数多，非标准读音常用在动词 hit，chew，swim 等之后，而标准读音常用在动词 criticize，visit 等之后，等一系列有趣的现象。由此总结出，把 ing 发成(ŋ)或(n)与说话人的性别、层次、个性、情绪以及谈话的正式程度和使用的动词等有关。费舍尔使用的方法比较简单，调查的人数也很少，可是他的研究方法非常值得我们借鉴。

(2) 建立假说

假设一般是根据已有的资料经过周密考虑后提出来，用以解释或说明客观现象之间的内在联系。假设一般有两类变量构成：一类叫自变量，起影响因素的作用；一类叫因变量，也就是被影响的因素。自变量和因变量的确立取决于研究目的。如在研究说话人的社会属性与其所用语言的特征之间的关系时，通常前者是自变量，后者是因变量。但也有人划定某种语言变体的使用范围着手去追踪说话人的社会特征，这样说话人的社会特征就成了因变量，而语言变体则成了自变量。两种变量之间存在着相关关系，可以根据相互消长分成正相关和负相关两种。

假设的建立必须注意以下两点：

第一，假设要具体可测。对假设的研究，结论要归结到假设的验证上去。拉波夫根据已有的调查表明在纽约人的心目中元音后的辅音 r 是一个非常敏感的语言变项，它有两种变异形式：发 r 音的是标准体，不发 r 音的是非标准体。他根据初步印象，提出了一个总的假设：假若纽约市本地人中有两个集团在社会分层的阶梯上处于高低不同的地位，那么他们在发 r 音上也会表现出相应的体现身份高低的差异。考虑到大百货公司为了使顾客满意，在语言表现上会力求达到顾客的标准，他进而提出了一个具体假设：高档百货公司售货员的话语中表现出 r 变量的最高值，中档百货公司售货员的话语中表现出 r 变量的中等值，而低等百货公司售货员的话语中表现出的 r 变量的值最低，这一假设是对调查研究可能获得的结果所作的预测。拉波夫在这一研究中并没有用统计的方法验证假设，只是对不同类型的人

群使用 r 音标准体的频率进行了简单的统计和比较。但是由于他的假设十分明确,调查的方法和过程非常清楚,资料的分析翔实有据,其结论也是很有说服力的。

第二,对自变量要加以限定。现实生活中影响自变量的因素是多种多样的,如果不注意限定自变量的范围,就会导致对观察的结果难以作出准确的评估和推断。

(3) 确定调查单位

调查目的确定后就要确定调查单位。最理想的调查是对被调查对象的全体进行逐一调查,但这样涉及的个体太多,比较理想的办法是进行抽样。

(4) 调查方法

调查方法一般可分为抽样调查和非抽样调查两种。

1) 抽样调查。所谓抽样调查就是指按照一定的计算方法,从研究对象的全体中根据随机原则,抽取部分单位作为研究的样本,通过对样本中个案的研究,并以此部分单位的指标数值去推断相关整体的指标数值。抽样调查的组织形式主要有以下几种。

① 简单随机抽样,又叫纯随机抽样。具体的组织形式是:调查者要把调查范围内的所有成员进行编号,逐个写在签条上,然后把签条放在一个容器内,充分搅和后,从中任意抽取若干个签条,签条上的号码所代表的单位组成的抽样总体就是样本。在抽取过程中,抽出签后,记下号码,然后将"签"重新放回,以保证每张签都有确定数量的抽签机会。如果出现了重码的现象,就应将该号码放回,重新抽取。简单随机抽样有一定的局限性:由于是随机抽样,所以有可能造成样本分布不均的问题。调查者在调查前,要事先对调查区域的人口等情况进行了解,否则会造成样本少,丧失代表性等问题。

② 系统抽样,也叫等距抽样。具体组织形式是:抽样时采用列有随意号码的表格,给每一个相关者一个号码。然后,调查者根据随意号码表,按数字顺序依次进行抽样。当每一个有指定号码者的号码同随意号码表中所选定的那个号码相符时,就被选定为取样对象,这样依次选取,直至调查者获得全部的被调查对象为止。除使用随意号码表外,调查者还可以在其选定的区域内,指定某个任意数作为他的取样对象,如每隔 5 个作为其取样的对象。取样的多少决定于取样人在取样人的总数中所占的比例。这种方法也有一定的局限性:它常常选进一些语言不能采用的人,也就是这些人的语

言是不能作为调查资料来进行研究的。

③ 分层抽样,又叫类型抽样。具体组织形式是:将总体单位按其属性特征分为若干层组或类型,然后从各层组或类型中随机抽取一定单位构成样本。分层抽样一定要做到界限分明,层组内的单位保持同质,层组间的成员要保持异质。各组的单位数可按定额分配,即不论各层组或类型的情况如何,一律抽取同样数量的单位,这种抽样方法也叫定额抽样。更为合理的办法是按比例分层抽样,即按层组或类型的总体单位数在全体单位数中所占的比例大小,分配样本数量。分层抽样适合在总体各单位标志值大小悬殊的情况下采用。运用分层抽样可以得到比简单随机抽样较为准确的结果。

④ 整群抽样。组织形式是将总体各单位划分成若干群,然后以群为单位从中随机抽取一些群,对所选取的群的所有成员进行调查研究。整群抽样的优点是组织起来比较方便,确定一群可以取得许多单位进行观察,但也会影响到全体及总体中各单位分配的均匀性,从而导致抽样误差大、代表性低等缺点。

⑤ 阶段抽样。当总体很大时,直接抽选总体单位在技术上是有很大困难的,因而一般采用分阶段进行的抽样调查组织形式,这种形式称为阶段抽样。阶段抽样适合用于大规模的复杂的社会方言调查。

抽样调查是一种非常科学的调查方法,是按照一定的计算方法和程序进行抽取调查单位的方法。但也有一些现实问题,即用计量方法调查研究社会方言,样本多大才算合理呢?这一般要根据调查要求而定。从理论上讲,样本越大,结论也就越可靠。

2) 非抽样调查。非随机抽样指的是按照调查人员的主观意图或取样的方便,从总体中抽取一部分单位作为调查对象。这种调查方法无法计算抽样误差和可信度,因而不能根据样本资料去推断总体情况。常用的非抽样调查方法有以下几类:

① 偶然性抽样。偶然性抽样就是指调查者在特定场合下以偶然遇到的对象作为样本进行调查的方法。这种调查方法偶然性大,材料不可靠,得出的结论一般也很难令人信服,一般不能作为正式的数据或依据。

② 雪球式抽样。这种方式首先选择一些调查单位,然后再从第一批调查单位的关系网中抽出第二批调查单位,依此类推下去,逐步扩大调查范围,最后得出全面、系统的调查结论。

③ 判断抽样。判断抽样一般是由研究人员自己去鉴定调查单位的代表性，然后从这些有代表性的典型调查单位中抽出样本来。

## 二、资料收集

语言调查所需收集的资料涉及三方面的内容：第一是语言使用者的资料，也就是说话人的身份、民族、性别、籍贯、年龄、职业、教育程度等等有关被调查者个人的背景情况；第二是使用语言的资料，也就是说话人在一定语境中所说的话语；第三是有关语言态度的资料，语言态度也就是言语共同体成员对特定的语言变体所持有的态度，包括看法、情感和行为意向。

### 1. 调查前的准备工作

(1) 制定调查大纲

调查大纲对调查者来说就是调查工作的计划。调查大纲是根据调查目的、任务而制定的。调查大纲一经制定，工作开展起来就能做到有条不紊，并能保证按时完成调查任务。调查大纲要有一定的预见性，在调查的过程中还要根据具体的主客观条件作出适当的完善和补充。一般的调查大纲要包括调查的目的、调查的项目、人员的分工、调查的步骤和时间安排、必须携带的调查表格、经费预算等等。

(2) 准备调查表格和相关语言材料

根据调查目的准备相关的调查表格和语言材料是调查前的一项必要的工作。譬如要调查汉语某一地区方言点的音系，那么就要准备《方言调查字表》和《汉语方言调查简表》等；若是进行词汇或语法调查就要准备《方言调查词汇手册》、《方言调查词汇表》和《现代汉语词典》等调查表格和工具书。

(3) 选择被调查者

被调查者的选择往往是通过中间人的帮助确定的，所以事先必须告诉中间人准备调查的方言点、人数、年龄、文化程度、合作时间等。理想的合作人应该是：口齿清楚、方音纯正、具有一定的文化程度、有良好的协作精神等等。

### 2. 收集言语资料和态度资料的方法

收集言语资料和态度资料的方法主要有4种途径，它们分别是采访被调查者；发放问卷；实地观察；进行试验。这只是就具体性质而言分出的相对独立的方法，在实际调查过程中这些方法往往是几种方法同时运用。

(1) 访谈法

访谈法就是由调查人员与被调查人员进行面对面的谈话，这又可以分成个别交谈法和集体交谈法两种。访谈法的关键在于如何引导受访者说出调查者所需要的话语，提供有关语言变体的第一手材料。为取得预期的效果，在调查过程中往往需要使用一系列技巧诱发被调查者。

1）念词表或句子。这是调查方言语音用得最多的技巧，调查人员把事先编制好的词汇表、例句或短文请被调查者逐一朗读，当场审音、记音。条件允许的情况下最好能够把访谈的内容录下来，以便访谈后仔细分析核实。

2）提问。这种方法是由调查人员向被调查人员提出问题，被调查人员按照调查人员的问题逐一进行回答，以便使调查人员从中获取所需要的信息。提问的方式可以是直接的，也可以是间接的。

直接的提问方式就是按照事先准备好的调查提纲，把所需了解的语言变异形式逐项提出来，询问被调查者平时经常说还是不经常说某种形式，或者是某种形式是否可以说。在提问的过程中，调查者应当尽量避免干扰，例如不要反问、不要诱导等等。直接提问法的缺点是被调查者的反应有时靠不住。如有一位家庭妇女在访谈中一再表明她只说“我们”，从来不说“姆末”，但采访结束送采访人员出门时，却一连说了5个“姆末”。

间接式提问方式即调查人员不直接询问有关语言的问题，而是提出一个话题，让被调查者讲述。值得注意的是这种提问方式选择的话题必须是被调查者能回答而且是感兴趣的问题。在交谈过程中调查者要及时引导，使被调查者能够说出所需要的东西。

3）测验。测验的作用在于给被调查者某种刺激，使其立即作出言语反应，以观察被调查者的语言能力，了解其语言体系的特定规律。测验的方法有很多种，如看图说话、变化句型、填空等等。访谈法的主要弊端在于被调查者意识到采访目的后，回答问题时会变得很谨慎、小心，影响调查效果。

（2）观察法

观察法就是调查人员通过观察研究对象的言语行为以获取资料的方法。观察法按方式可以分成以下几类：

1）隐蔽观察法。隐蔽观察法就是观察人员在人们不注意的情况下观察他们的言语行为。由于调查的方式比较隐蔽，被观察者不会受到其他因素的影响，调查到的是比较自然的言语状态，缺点是无法控制变量，对说话人的社会特征难以搞清楚。这种调查比较适于探索性调查，了解一般的语言状况。

2) 迅速隐匿观察法。迅速隐匿观察法就是调查者采用巧妙的诱导技术,在不露真相的情况下让被观察者说出所需要了解的语言变异形式,迅速收集到所需言语材料。这种方法是拉波夫在调查纽约百货公司的社会分层时首先使用的。

3) 参与观察法。参与调查法就是指调查者参与到被调查者中去,与他们一起参与言语活动,在其中观察和收集资料。这种调查方法调查者可以不暴露自己的身份,或者是暴露自己的身份但不暴露自己的目的。这种方法的优点是:调查者和被调查者打成一片,不仅能收集到大量的自然言语材料,而且还能了解到被调查者的社会情况。

语言调查总会遇到一个方法上的困难,有人叫做观察者的矛盾。收集资料的可靠方法是进行个别谈话,把谈话的内容录下来。但是在正式访谈时,被访者往往因为意识到采访者的目的而有意地或不自觉地调整自己的语体,说出的话比平时正规,这就会影响到调查者收集到的资料的可靠性。所以有时可以综合运用 3 种方法,以期得到最自然的言语材料。

(3) 问卷法

问卷法,又叫填表法,是采用书面的形式进行语言调查的一种方法。问卷法是先要根据调查目的设计出一定数量的问题(包括被调查者的个人信息),编制成问卷,然后要求被调查者对问题逐一作出解答。问卷调查的方式可分为两种:一种是封闭式的,即把问题的若干种可能的答案用简洁的语言表达清楚,供被调查者选择;一种是开放式的,即对被调查者的回答不限定范围,由被调查者任意回答。

问卷调查适宜于大规模的摸底调查。它的问题是对某些个别被调查者,或对某些小群体的直接观察,并不总是具有足够的代表性,因为往往很难排除他们所具有的特征,很难把规范现象和某些个人偏离规范的现象区别开来。所以在设计调查问卷时,有时需要先作一些必要的探索性的调查。

(4) 试验法

试验法是研究人员有计划地运用某种手段控制、干预或模拟语言现象,以取得研究资料的方法。这种方法主要用于测量人们对语言的态度和反应。采用实验法要注意以下几点。

1) 排除其他干扰,控制自变量。最简便的了解人们对语言变体的态度和评价的方法,是直接向被调查者询问他的看法。但是其缺点是被调查人

的回答有可能不真实，调查人员对这种回答很难作出精确的分析。也有人采用一种被称作是主观反应测验的办法，它主要是在被测验人不知道进行测验的目的的情况下，把不同的语言变体话语或朗读录音的材料放给他听，听完后请他对有关问卷上说话人的态度作出评判。说话人的态度主要包括人的价值观念、职位的高低、性格特点、智力情况等等。主观反应测验法也存在一个问题，就是被测验人的性别、音色等因素可能会影响到测验的效果。因此要想获得对某种语言的确切评价，就必须控制住其他的自变量，排除非主要成分的干扰。

2）进行事前和事后测验。事前事后测验的作用是通过对比事前测验和事后测验的结果，了解实验本身是否改变了被测验者的态度。

3）设置控制组。实验的基本内容是观察和测试自变量对因变量的影响。但有时实验本身也会对因变量产生影响，实验是否影响了因变量可以通过设置控制组的办法进行鉴别。具体做法是：事前测验以后，把被测验人分成实验组和控制组两组，实验组的成员接受测验，控制组的成员不进行测验，最后两组同时进行事后测验，然后对两组的测验结果进行比较，如果双方事后测验结果与事前测验结果没有明显的变化，说明测验本身没有对被测验者产生什么影响，如果事前与事后的实验结果有较大变化，说明测验对被测验者产生了影响。而且如果测验结果相差非常大，就要考虑是否还有其他因素影响了测验。

## 三、资料整理

资料收集完之后就要对它们进行分类和汇总。资料的整理和分析主要采用两种方法，一是定性分析，一是定量分析。定性分析就是确定调查所获得的资料是否适用于调查者的调查目的，而定量分析要远比定性分析复杂得多，所以下面我们详细介绍。

1．资料量化

1）定类性。适用于量度层次分明的研究对象——社会变量或语言变量。调查者对收集的与调查目的有关的资料进行的首要工作就是对资料的分类，并给不同类型的资料加上不同的标注，以示区别。两个相邻的变量值之间没有中间数值的叫离散型变量。

2）定序性。有些变量之间不是有无的对立，其各形式间是连续不断的，相邻的两个值之间可作无限的切分，这种变量被称作连续变量。对于连

续型变量我们只能采取定序的方法反映相互间的高低、大小、强弱等一系列等级差别。定序性有时可以按照调查者的目的转化成定类性。

2. 资料统计

资料的定性定量工作完成后，就需进行资料的统计工作，常用的方法有：平均值、标准差、概率和正态分布 4 种。

1）平均值。平均值是指样本内的各个变量值所达到的平均水平，它可以反映分布的集中趋势。平均值的计算公式是：

$Y=\sum X/n$（$Y$ 为平均值，$\sum$ 表示总合，$X$ 为变量，$n$ 为单位总量）

2）标准差。标准差是用以测定离中趋势的标志变异指标。离中趋势是指在数列中各变量值的离散度或变异度。要说明各变量值的分布状况，必须把平均值与标准差结合起来运用，因为平均值相等的两组变量值，其标准差未必一致。标准差和平均值之间的关系为：标准差越大，平均值的代表性越小；标准差越小，平均值的代表性越大。标准差有两种，一种是样本标准差，一个是总体标准差。样本标准差的计算公式是：

$$S=\sqrt{\frac{\sum x^2-\left[\left(\sum x\right)^2/n\right]}{n-1}}$$

标准差的平方叫做方差，样本方差用 $S^2$ 表示，其计算公式是：

$$S^2=\frac{\sum x^2-\left(\sum x\right)^2/n}{n-1}$$

在掌握了全部资料后，就可以求得总体标准差的值，其计算公式是：

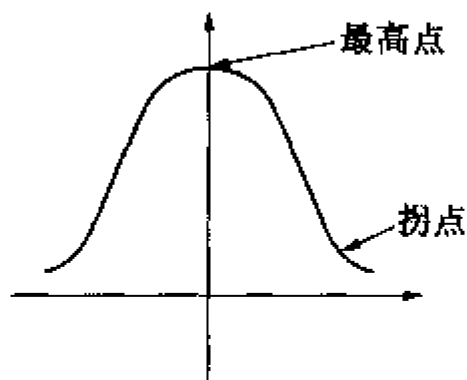
$$\sigma=\sqrt{\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{N}}$$

3）概率。概率是用来描写某种现象出现可能性的大小的数量表达。它可以告诉我们在一次观察或实验中出现的偶然现象，在相同条件下通过无限量的重复观察或实验，会呈现出一定的规律性。概率的统计定义是：在一组不变的条件下，重复做几次试验，$\mu$ 是 $n$ 次试验中事件 A 出现的次数。如果当试验次数 $n$ 足够大时，事件 A 出现的概率（$=\mu/n$）稳定地在某一数值 $P$ 附近摆动，并随着试验次数的增加，越来越接近 $P$，那么 A 称为随机事

件，$P$ 就称为事件 A 出现的频率。

4）正态分布。正态分布是概率的常态分布，用曲线图表示呈钟形。

正态曲线图取决于两个参数，一个是平均值，一个是方差。平均值和方差的不同组合对应于不同的正态曲线，纵轴将正态曲线分成相等的两部分，正态曲线图在其平均值处达到最高点，在平均值＋1 个标准差处是拐点。从拐点向外延伸逐渐接近横轴，但不相交。




## 参考书目

1. 冯志伟. 应用语言学综论. 广东教育出版社，1999

2. 冯志伟. 中文信息处理与汉语研究. 商务印书馆，1992

3. 吕冀平. 当前我国语言文字的规范化问题. 上海教育出版社，2000

4. 苏培成. 现代汉字学纲要. 北京大学出版社，1994

5. 许嘉璐、王福祥、刘润清. 中国语言学现状与展望. 外语教学与研究出版社，1996

6. 于根元. 语言应用论集. 北京广播学院出版社，1999

7. 真田信治、涩谷胜已、阵内正敬、杉户清树. 社会语言学概论. 王素梅、彭国跃译. 上海译文出版社，2002

8. 祝畹瑾. 社会语言学概论. 湖南教育出版社，1992

9. A · Д · 什维策尔. 现代社会语言学. 卫志强译. 北京大学出版社，1987

10. R · A · 赫德森. 社会语言学. 丁信善译. 中国社会科学出版社，1990

11. 许余龙. 对比语言学概论. 上海外语教育出版社，1992

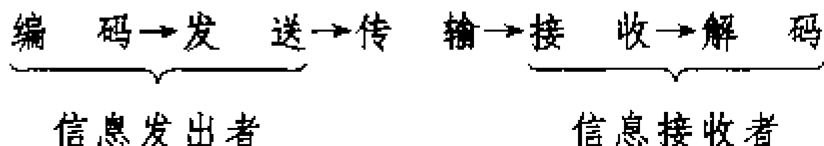
# 第六章 社会语言学

## 第一节 社会语言学的性质

### 一、语言与社会交际

1. 交际

人是社会关系的总和,就是说人与人之间总是存在于一定的社会关系之中。处于各种社会关系中的人不是孤立的个体,不是自然物,人与人之间需要联系,需要交往和沟通。人与人的交往和沟通就是交际。所谓交际(communication)就是利用某种已经建立起来的公认的符号系统来有意识地传递信息,从而起到交流思想、传达情感、协调关系的目的。交际是一个动态的过程,如果借用信息论的观点,一个最简单的交际过程可以概括如下:




信息发出者和信息接收者是交际双方,信息发出者想要传递信息,必须用交际双方都能理解的一套符号系统来承载信息,即对信息进行编码,编码完成后要将信息发送出去,信息发送后就进入了传输过程,即进入信道(channel),原始的传输信道是空气,后来有了文字媒介,再后来有了电话、电视、电报、传真、电子邮件、网络等电子媒介。对于信息接收者来说,他接收到信息后,要进行解码,然后才能理解。信息接收者,无论理解了还是没有理解,都可能作出反馈。

在整个信息传递过程中,承载信息的符号系统极其重要。交际双方都需要一套传递信息的符号,人类所使用的传递信息的符号可以分两大类,一

是语言符号，二是非语言符号。非语言符号包括体态语(body language)、副语言(paralanguage，包括沉默、话轮转换、各种非语义声音)、客体语(object language，主要指交际者自身的肤色、着装、化妆、居室等信息)、环境语(environmental language，包括空间信息、时间信息、声光、颜色、标识符号等)。不过，人与人交际时所使用的最主要、最复杂，也是最轻便的符号系统是语言符号。语言是人类最重要的交际工具。

语言交际是一个复杂的动态过程，交际是否能顺利进行涉及许多因素，如交际的主体和客体(听说双方)，交际的目的和主旨，交际的时间和空间，交际的社会背景，等等。反过来说，这些因素都影响了语言交际，也影响了交际时对语言的选择和运用。交际是具有社会性的，用来交际的工具——语言，也必然是具有社会性的，社会性是语言的本质属性。

**2. 语言与社会的关系**

作为人类社会交际工具的语言跟社会有着密切的关系。这种关系可以从如下几个方面来观察。

(1) 社会的存在和发展离不开语言

所谓“社会”就是指由一定的经济基础和上层建筑构成的整体，也指一定的“人群”或人类共同体，如部落、部族和民族，他们生活在一个共同的地域中，说同一种语言，有共同的风俗习惯和文化传统。同一社会中的人需要共同劳动、共同生活、共同斗争，就需要有帮助沟通、交流、协调的工具，这个“工具”就是语言，没有语言人与人之间的联系就会中断，社会就会解体。

社会的发展和进步也离不开语言，社会发展中所积累的经验和教训，需要通过语言来记录和传承，推动社会发展和进步的政治、法律、制度、文化、科技需要语言的记录，需要通过语言来传播，新的发现和发明需要通过语言来记载和公布，人类的未来在于教育，教育更是离不开语言这个工具。

(2) 语言的产生和存在离不开社会

语言随着社会的产生而产生，没有人类社会就没有人类的语言，人类之外是没有语言的，语言只不过是人类社会的交际工具。如果某个部落、部族或民族在历史发展进程中被别的民族融合了，或者因为其他原因而被灭亡了，该部落、部族或民族的语言也就不存在了。如美洲印第安语原有许许多多土著语言，但随着战争、屠杀、迁徙、瘟疫、同化，许多土著部落或部族消亡了，其部落或部族语言也就消亡了，世界上许多土著语言正随着土著部落或部族的消亡而消亡。再如，我国历史也曾经发生过多次民族的融合和语言

的融合。如从春秋战国起，史书上就记载有“东夷”、“南蛮”、“西戎”、“北狄”等民族，即在汉族周围居住着许多兄弟民族，这些民族都有自己的语言。《左传·襄公十四年》记载戎子驹支的话说：“我诸戎饮食衣服，不与华同，贽币不通，言语不达。”另外，刘向《说苑·善说》记载，当时越人唱的“越人歌”，楚国人子晰竟然听不懂，需要翻译。这些记载都说明，当时汉族周围的一些民族的语言跟汉语是不同的，相互间不能通话。可是，经过春秋战国时期的会盟、战争、兼并、同化等，发生了民族的融合和语言的融合，结果是在后来的史书中再也看不到关于这些民族和语言的记载。两汉以后，北方的匈奴、鲜卑、羌、羯以及后来的契丹等民族也在同汉族的融合中失去了民族的独立性，其语言也在同汉语的融合中被同化。可见，语言的存在离不开社会，社会的消亡必然带来语言的消亡。

不仅如此，语言随着社会的变化而变化。社会由低级到高级，由简单到复杂，由落后到先进的发展，都会推动语言的发展；社会的分化、统一、相互接触也会引起语言的分化、统一和接触。就某种具体语言来说，随着社会的发展，如新事物的产生、旧事物的消亡、人的思维的发展，都会带来词汇的变化和语法的改进。现代汉语跟古代汉语相比，词汇大大地丰富了，语法更加趋于缜密了。尤其是社会的急剧变化，在语言中的反映更加明显，在汉语发展史上，春秋战国时期、魏晋南北朝、晚清和五四时期、1949 年前后、近 20 年中，不仅是社会急剧变革的时期，也是汉语发展和变化最显著的时期。

(3) 语言规划和语言政策是由社会决定的

语言的发展和变化有语言内部因素的制约，也有语言外部因素的制约。随着人类社会的发展，许多国家和民族对语言的发展和运用进行了政策性的或人为的干预，这种干预主要体现在语言规划和语言政策的制定上。而语言政策的制定不是由个人决定的，而是由具体的社会因素决定的。如标准语的选定和规范是由社会决定的，某种方言能否成为标准语的基础方言，是由该方言所处之地区的政治、经济、文化的地位决定的，政治、经济、文化中心区的方言往往能成为基础方言，如现代汉语普通话选择北方方言为基础方言，就是因为北方话区，尤其是北京，相当长时期内是中国的政治、经济、文化的中心，北京话有着广泛的影响。再如，双语和多语国家的官方语言的确定，也要考虑到政治、经济、文化、民族关系等社会因素。

(4) 语言行为同时也是社会行为

作为社会的人其所能使用的语言或方言往往是多样的，有的人只会一

种方言，有的人不仅会一种方言，而且会说标准语，有的人则既能说标准语，也会说多个方言，有的人能操双语或多语。一个能说多种语言或方言的人在具体交际中就出现语言或方言的选择问题，这种选择往往不是说话者的个人爱好，而是由社会决定的，即语言对社会的认同，或语言社团的归属感。因为语言是一个民族或社会的灵魂，当一个人在使用一种语言时，就是在表明他是这个民族或社会的一分子。当双语或双方言（多语或多方言）人士在选择某种语言或方言进行交际时就表明了他对该语言的态度，或者是他认为在这个确定的场合选择该语言或方言是得体的，或者他认为选择该语言或方言更能表达他的身份和社团归属。双语或多语的地区或国家的语言冲突实质上是民族或社会团体的冲突在语言上的表现[①]。即使在同一种语言或方言内，语言形式的各种变异也跟社会经济等级或"社会阶级"有关[②]。可见，语言行为实质上同时是社会行为。

总之，语言和社会具有共变关系，语言离不开社会，社会不能没有语言，社会促进了语言的发展，语言反映了社会的发展。正因为语言跟社会有着如此密切的关系，研究语言与社会的关系的社会语言学就是十分必要的，也是必然的。

## 二、社会语言学

### 1. 社会语言学的产生

联系社会来研究语言是很早就有了，如我国古代哲学家荀子就正确地论证了语词音义结合的社会约定俗成性："名无固宜，约之以命，约定俗成谓之宜，异于约者谓之不宜。"秦始皇时代就开始有意识地进行"语言规划"——书同文。在古代字书、辞书如《说文解字》、《尔雅》中处处可以见到作者联系社会来解释字词的现象。汉代扬雄的《方言》（輶轩使者绝代语释

① 如加拿大宪法规定，英语和法语都是加拿大的官方语言，但魁北克省是加拿大法语集中地，法语意识特别强烈，1977年魁北克省议会通过了著名的《101号法律》，规定法语为该省的惟一正式语言，禁止使用其他语言，但这引起了该省讲英语的人的不满和抗议，于是议会又通过了《198号法律》，规定室外广告只允许使用法语，室内广告可以使用其他语言，这才使矛盾暂时缓和。语言的地位在加拿大是一个敏感的社会问题。

② 如拉波夫（Labov）关于纽约市百货公司（r）的社会分层研究（Sociolinguistic Patterns, Philadelphia: University of Pennsylvania Press）。拉波夫把纽约市民按照/r/元音发音情况分为6个阶层：上中等、下中等、上层工人、中层工人、下层工人、低层。社会经济地位越高的阶层，讲话时卷舌音的比率就越高。

别国方言)就注意了对语言的地域变体——方言的调查。这种传统历代不绝。尤其到了20世纪四五十年代罗常培更是明确注意到要联系社会、文化、历史来研究语言,并在50年代出版了专著《语言与文化》,这是一部系统地论述语言与社会文化关系的著作。在西方,早在18、19世纪,德国学者如赫尔德、洪堡特等就论述了语言和民族的关系,19世纪末20世纪初,法国的社会学派如梅耶等就将语言研究的社会方面纳入语言学领域,20世纪初,英国人类学者马林诺夫斯基呼吁建立"人类文化语言学",马林诺夫斯基以及弗斯为代表的伦敦学派其基本语言观就是社会语言学的,重视语言使用环境对语言的影响。不过,20世纪60年代以前的研究,是零散的,不成系统的,社会语言学作为语言学的一个分支学科并没有建立。

社会语言学作为语言学的一个分支学科,产生于20世纪60年代中期。1964年5月,由语言学家和社会学家发起在美国召开了首届社会语言学大会,会后由布赖特编辑了论文集《社会语言学》,并于1966年出版,论文集提出了社会语言学研究的诸方面问题。会议和论文集所包括的内容,说明社会语言学(sociolinguistics)已经被正式提出①。60年代以后社会语言学受到语言学界的广泛重视,不仅有相当多的社会语言学专家,有大量的论著,而且还有专门的学术杂志,如 Language of Society(《语言中的社会》)、International Journal of the Sociology of Language(《国际语言社会学杂志》)以及 Language Variation and Change(《语言的变异和变化》)、Journal of sociolinguistics(《社会语言学杂志》)等。

社会语言学的诞生是语言学发展到一定阶段的产物。传统语言学跟社会文化有着密切的关系,历史比较语言学也离不开对社会历史的分析,但这些研究都忽视了语言单位、语言要素之间的互相依赖、相互制约的关系,而把语言事实作为孤立的现象来处理。20世纪初出现的结构主义语言学、转换生成语言学强调从语言结构内部来研究语言,把语言作为一个独立的系统来处理,把探讨语言内部规律作为惟一的任务,忽视了从语言外部来研究语言。60年代以后,许多学者认识到语言并不是一个独立的、封闭的系统,于是产生了许多新的语言学分支学科,如心理语言学、神经语言学、语用学、计算语言学、人类语言学、文化语言学等,分别从语言外部的不同角度来研

---

① 有人认为社会语言学这一术语最早由美国语言学家丘里在1952年的论文《社会语言学的设计:语言和社会阶层的关系》中提出。

究语言，社会语言学就是这些新学科的一种。

另外，社会语言学的兴起，还跟相关社会科学如社会学、民族学、人类学、心理学、文化学的成熟和发展有密切的关系，这些学科为社会语言学的形成和发展提供了许多有用的概念和方法。

2. **社会语言学的性质和特点**

关于社会语言学的性质有不同的看法，大致说来有 3 种认识，一种观点认为社会语言学是联系社会研究语言的一门学科，强调社会语言学是语言学的一个分支学科；另一种观点认为该学科是联系语言来研究社会的，即用语言材料来描写和解释社会行为，这种观点被称为语言的社会学或语言社会学(sociology of language)，这种观点实际上是倾向于把社会语言学看作社会学的范畴；再一种观点是强调语言跟社会的共变关系，既可以从社会看语言，也可以从语言看社会，社会语言学正好处于社会学和语言学中间，这实际上是持一种实用的观点。我们认为这 3 种观点表明了研究者的不同侧重点，实际上不同观点的研究对象大体上是一致的，即语言和社会。

我们认为社会语言学是属于语言学的，它是运用语言学和社会学等学科的理论和方法，联系社会来研究语言现象的一门学科，主要是研究语言的社会属性和差异。社会语言学虽然只有不太长的发展历史，但已经形成了自身的特点。

首先社会语言学具有综合性。社会语言学要借助于其他学科的理论和方法来观察和研究语言现象和语言问题，社会语言学的建立就是语言学家、社会学家、人类学家等共同努力的结果，研究社会语言学的问题既要语言学知识，也需要社会学、人类学、文化学、民族学、教育学的知识。如研究语言的社会分化(社会方言)，要涉及社会学中关于社会分层、社团确立等社会学观念，研究语言的各种变异的具体情况需借助于社会学的统计、调查等方法；研究双语政策、双语教育、语言教育规划，就需要教育学的相关知识；研究语言跟民族、文化的关系就需要民族学、人类学、文化学的相关知识；研究语言的地域变体就跟地理学、移民史等有关。总之，社会语言学是一门综合性、边缘性极强，具有跨学科性质的学科，研究社会语言学也要寻找好语言学跟其他学科的结合点，充分发挥社会语言学的综合性优势。

社会语言学具有很强的应用性。社会语言学的成果可以为语言规划、语言教育、双语教育等提供政策性依据，社会语言学的应用在语言文字的地位规划和本体规划上的表现尤为突出，社会语言学可以为双语教育态度的

分析、双语教育模式的制订或选择、双语教育的评估等提供理论上的依据；语言变异跟社会阶层之间的对应关系研究成果可以为语言社区教育（如语言偏见矫正）提供指导；语言跟文化关系研究可以应用于跨文化交际；语言和社会关系的研究可以为词语的社会意义研究提供线索，像"非婚子女"、"单亲家庭"、"婚内强奸"、"无罪推断"、"水泥枕木"、"第三者"、"二奶"、"丁克家庭"、"网络化生存"、"数字化生存"等词语的意义，如果脱离社会语境是很难解释和理解的。另外，社会语言学研究成果还在商业（如广告文化、商品品牌名禁忌）、法律（如法律语言的话语分析、功能辨别，法律文本的可读性指导，法庭背景下的文化误度的消除）、医学（如医生跟病人交谈时的话语选择、交谈方式、话语模式等）、行政文书（如语体和用词）、第二语言习得（如交际适应理论和中介语问题，涉及跨文化交际）、文学研究（如文学话语分析模式中的父权主义分析、女权主义分析、社会语言学分析等就是社会语言学为文学研究所提供的一种新的观察视野，社会语言学还可以为文学作品中人物角色的社会地位、身份认同、叙事方式、话语模式、语言表达等的分析提供帮助）等方面有较为明显的实用价值。从这个角度看，社会语言学属于语言学中的应用语言学的一个分支学科。

社会语言学具有实验性。由于社会语言学是联系社会来研究语言问题和语言现象的，因而社会语言学需要借助社会调查法、科学实验法和现代统计测量的方法。像拉波夫的代表作《纽约市英语的社会分层》(1966)这篇社会语言学的经典论著就采用了社会调查和科学实验的方法，他用随机抽样的方法来抽取和选择调查单位，用社会变量值作为划分调查对象层次的尺度，又用谈话法和观察法设法诱导和分离出随情景而异的语体，用录音机记录调查素材，对调查得到的材料又用数学方法进行统计和量化分析，最终使语言变项跟社会阶层对应起来，从而使社会语言学取得了突破性的进展，并建立了社会语言学中的"城市方言学派"。

最后，要强调的是社会语言学的语言学性质。社会语言学虽然重视社会跟语言的共变关系，研究语言和社会的各种关系，但它本质上是属于语言学的，只不过是强调从社会角度来探讨语言的社会本质和语言变异的社会原因。

3. 社会语言学的任务

社会语言学是一门年轻的学科，其研究对象不同的人看法不尽相同，但概括起来看不外乎以下一些方面：① 社会因素如民族、行业、阶层、环境、地域等对语言结构的影响；② 社会因素对语言功能、语言地位的制约作

用；③ 与语言使用有关的宏观问题，如语言政策、语言规划、语言规范、语言教学、双语(双方言)现象、双语教育等；④ 由社会因素引起的语言之间的关系，如语言接触、语言影响、语言融合、混合语等；⑤ 个人因素对语言的影响，个人因素如职业、年龄、性别、文化、心理、性格、信仰等；⑥ 社会因素在语言上的种种反映，从语言变异看社会变异；⑦ 语言因素对社会因素的影响；⑧ 人们怎样在实际环境中使用语言进行交际和不同社会、社团使用语言的差别。

### 三、社会语言学研究和流派

#### 1. 国外社会语言学研究

社会语言学诞生于20世纪60年代的美国，其后美国的社会语言学主要分两个流派，一是以拉波夫为代表的城市方言学派，也称为微观社会语言学派，拉波夫的代表作有《社会语言学模式》(1972)，另一个是以费什曼为代表的所谓宏观社会语言学派，其代表作是《社会语言学初阶》(1972)。另外，美国学者 Ralph Fasold 还出版了《社会的社会语言学》(1984)和《语言的社会语言学》(1990)，这两部著作分别从不同角度论述了社会语言学的中心问题，是社会语言学的权威著作。此外被翻译到中国的美国学者的著作还有佐伊基《社会语言学演讲录》(1989)等。

英国学者彼德·特鲁杰尔《社会语言学导论》(1972，中译本由陕西人民出版社1990年出版)、赫德森《社会语言学》(1981年，中译本由中国社会科学出版社1990年出版)也是较有影响的通论性著作，这两本书因为有中译本，在中国较有影响。前苏联和东欧学者在社会语言学研究上也有不少贡献，如前苏联学者什维策尔的《现代社会语言学》(1987)，前南斯拉夫学者 Milorad Radovanovic 的《社会语言学》也有一定影响。日本学者在社会语言学研究上也有不少成果，如柴田武的《社会语言学课题》(1978)、松山千秀和生田少子的《社会语言学》(1988)、真田信治的《社会语言学》(1992，中译本，上海译文出版社，2002)、真田信治的《日本的社会语言学》(1994，中译本，中国书籍出版社，1996)等。

#### 2. 中国的社会语言学研究

20世纪50年代罗常培的《语言与文化》应该是中国较早的社会语言学研究专著。但中国对社会语言学的真正研究开始于20世纪的70年代末。70年代末和80年代初，《语言学动态》和《国外语言学》等杂志介绍了国外

社会语言学的一些研究成果，这些译介推动了中国社会语言学的研究。1979 年和 1980 年陈原的《语言与社会生活——社会语言学札记》分别在香港和内地出版(三联书店)，其后陈原又出版了《社会语言学》(学林出版社，1983)、《社会语言学专题四讲》(语文出版社，1988)、《社会语言学论丛》(湖南出版社，1991)、《语言和人——应用社会语言学若干探索》(上海教育出版社，1992，商务印书馆，2003)等。陈原的研究对中国社会语言学的建立和发展具有开创价值。20 世纪 80 年代以后，中国的社会语言学研究发展迅速，大学里开始开设社会语言学课程，招收了一定数量的社会语言学研究生(语言学及应用语言学的社会语言学方向)，成立了专门的研究机构，如国家语委语言文字应用研究所社会语言学研究室、中国社会科学院少数民族语言研究中心社会语言学研究室等，召开了 3 次全国社会语言学学术研讨会和首届社会语言学国际学术研讨会，出版了一定数量的社会语言学专著，如陈松岑《社会语言学导论》(北京大学出版社，1985)、孙维张《汉语社会语言学》(贵州人民出版社，1991)、戴庆厦《社会语言学教程》(中央民族学院出版社，1993，中央民族大学出版社，1996)、申小龙《社区文化与语言变异——社会语言学纵横谈》(吉林教育出版社，1991)、王得杏《社会语言学导论》(北京语言学院出版社，1992)、祝畹瑾《社会语言学概论》(湖南教育出版社，1992)、徐大明等《当代社会语言学》(中国社会科学出版社，1997)、郭熙《中国社会语言学》(南京大学出版社，1999)、陈建民《中国语言和中国社会》(广东教育出版社，2001)等。近年来在社会语言学专题调查、描写和研究方面也出版了一些重要成果，如陈建民《语言文化社会新探》(上海教育出版社，1989)、陈建民和谭志明主编《语言与文化多学科研究》(北京语言学院出版社，1992)、张清常《胡同及其他——社会语言学探索》(北京语言学院出版社，1990)、陈松岑的《语言变异研究》(广东教育出版社，2001)等，其中双语双方言研究、文化语言学研究、语言接触研究、言语交际研究、民俗语言研究、民族语言和文化研究、网络语言研究等取得了更多的成果。

3. *社会语言学流派*

中外社会语言学都取得了许多成果，这些研究从取向上看，主要可以分为两个流派，一是微观社会语言学(micro-sociolinguistics)，或称小社会语言学，主要研究语言的各种变异(variation)，并且联系社会因素来探讨语言变异发生的原因和规律，常常是用统计的方法和概率的模式来描写这些变异现象，微观社会语言学的经常性课题是语言与社会、语言与阶级、语言与

环境、语言与性别、语言与种族集团等，实际上是讨论各种社会因素对语言变异的影响。微观社会语言学早期是以拉波夫的“城市社会方言学派”为代表，在中国，陈原、陈松岑、陈建民、张清常、祝畹瑾等的研究基本上属于微观社会语言学范围。另一个流派是宏观社会语言学(macro-sociolinguistics)，或称大社会语言学，主要研究社会中的语言问题，如双语、双方言或多语的交际与教育，语言政策，语言规划，标准语选择，语言的相互接触与影响，语言冲突等，这些问题往往由于社会因素而产生，也需要从社会的角度加以研究和解决。宏观社会语言学在 20 世纪 70 年代后期以费什曼为代表，主要从社会学角度研究语言问题。宏观社会语言学在中国历来受到重视，中国近百年来的语文现代化运动(如汉字改革、白话文运动、国语运动、拼音化运动、推广普通话、《汉语拼音方案》的制订、《中华人民共和国国家通用语言文字法》的制订、少数民族语言政策制订和少数民族文字的创制等)实质上都属于宏观社会语言学范畴。

4. 社会语言学研究方法

社会语言学的研究方法，总是综合采用语言学、社会学、人类学、民俗学、文化学、心理学、教育学、统计学等方法以及自然科学中的某些方法。具体来说有以下几种方法：

(1) 调查法

调查是社会语言学最重要最基本的研究方法。一般来说，社会语言学的调查有实地调查和通信调查(包括网上调查)，全面调查和抽样调查，口头调查和书面调查，还有录音录像调查等，比较常用的是通信调查和抽样调查。社会语言学尤其重视口语调查，尤其要记录好未经加工的原始口语。社会语言学调查要按照调查的目标设计好调查项目和调查表格，如语言变异的项目清单和可以涉及的社会因素等。

(2) 描写

描写是社会语言学的基本方法，社会语言学描写的是语言应用的各种变异情况以及变异的原因、规律和语言运用的层次分化状况等。

(3) 分类

主要是分清语言运用和变异中的相似点和不同点，在相似点中又要区分不同的情况和不同的层次，总结出社会因素对语言影响的不同类型。

(4) 分析和统计

分析有定量分析和定性分析，社会语言学更加重视定量分析，主要是对

调查所得的材料用人工或计算机(当前主要是利用计算机)进行数量分析和统计,得出科学的数据,具体说明语言运用中的各种变异情况等。

## 第二节 宏观社会语言学

宏观社会语言学主要指语言接触和语言规划,包括双语、双方言或多语的交际与教育,语言政策,语言规划,标准语选择,语言的相互接触与影响,语言冲突等问题。语言政策、语言规划、标准语选择等内容由于在汉语应用语言学中占重要地位,本书将在第五章中专门讨论。本章主要介绍跟语言接触有关的现象。

语言接触(language contact)是一个内涵广泛的概念,主要指不同民族之间由于社会、文化的相互接触和影响在语言上出现的种种关系,如语言影响、语言兼用、语言转用、语言混用、语言冲突等。

### 一、语言影响

语言影响(language influence)是指某一语言受另外语言的影响。语言的影响是经常性的,有的民族因为地域相连,不同民族语言之间自然相互影响,有的则可能因为翻译或第二语言学习而产生影响,有的可能因为科技文化的借鉴而产生语言的影响。语言影响可以从不同角度去观察。

**1. 从影响的范围上去观察**

语言影响涉及语音、词汇、语法等各个方面,其中以词汇受影响最为明显。不同语言互相接触,总是先出现词汇的借用,外来词或借词是语言的普遍现象。在汉语中,外来词由来已久,如“葡萄”、“石榴”、“菠萝”、“狮子”、“玻璃”等是汉代从西域借入的词,“佛”、“菩萨”、“罗汉”、“阎罗”、“和尚”等是汉代以后从印度借入的词,“胡同”、“站”(驿站、车站)、“蘑菇”等是元代从蒙古语中借入的词。而更多的外来词是鸦片战争以后,尤其是五四新文化运动以来,从西方语言和日语中借入的词,如“沙发”、“可可”、“雷达”、“尼龙”、“纳粹”、“伏特加”、“CT”、“卡拉 OK”、“VCD”、“E-mail”、“思想”、“具体”、“演绎”、“理性”、“间谍”、“卫生”等。不过汉语中借用外来词常用意译的方式,意译词不是真正的外来词,如“足球”、“电话”、“钢琴”、“面包”、“火

箭炮"、"黑板"、"铁路"等，所以，汉语中外来词的比重并不大。

有的语言，如英语，它的外来词的比重比较大，英语中一半左右的词语是外来词，其中又以来自法语的外来词为多。1066 年，法国诺曼王威廉击败了英吉利军队，在英国建立了王朝，法语成了国家、宫廷、教会的语言，就在这一时期，法语对英语产生了很大的影响，英语从法语中借入了大量的词。外来词被称为"异文化的使者"，是民族的接触在民族语言中的反映，也是民族接触的见证。

语言影响中，不仅词汇可以借用，语法规则和语音也可以借用。如中国的一些少数民族语言往往受汉语语法规则的影响，增加了新的语序，如壮语中名词修饰名词时的固有语序是中心语名词在前，修饰语名词在后(如：房草→草房)，由于受汉语中修饰语在前、中心语在后的语序的影响，壮语中增加了名词定语在名词中心语之前的新语序(如：中国人民)。再如，民家语原有的基本语序是：主语＋宾语＋谓语，受汉语的影响，现在大多数情况下都采用了汉语的基本语序：主语＋谓语＋宾语。一些少数民族语言受汉语的影响增加了新的音位，如哈尼语受汉语的影响增加了"f"声母，侗语中[ph]、[th]、[kh]等送气音原来只是不送气音的音位变体，现在受汉语的影响已经成为独立的音位。

2. 从影响程度上去观察

语言的影响有程度上的不同。较小的影响只涉及少量的借词，如汉语中只有少量的俄语借词；而汉语对境内一些少数民族语言的影响则较为深一些，不仅有较多的借词，而且有的少数民族语言增加了语序，有的增加了音位。更深的影响则可能引起语言类型的变化，如海南省三亚回族使用的回辉话的变化就属于这一类。回辉话的基础是属于南岛语系的一种语言，南岛语属于粘着语，以词根附加法为主要语法手段，常用词大多为双音节，没有声调。回辉话由于长期受汉藏语的影响，发生了类型的变化，由粘着型变为孤立型，常用词由双音节变为单音节，并出现了声调。

3. 从受什么语言的影响上去观察

语言的影响从受何种语言的影响上看，有的受主体民族语言的影响，有的则受邻近的少数民族语言的影响，这在多民族语言国家中都有表现。如在我国，民族的分布特点和发展状况是决定受何种语言影响的主要因素，一般来说，各少数民族语言总是受到汉语的影响，因为汉族人口多，分布广，又是各民族之间共同使用的交际语言和国家标准语。另外，使用人口多的少

数民族语言也会对其周边的少数民族语言产生影响，如藏语对门巴语、珞巴语的影响，维吾尔语对哈萨克语、乌兹别克语的影响，傣语对景颇语、阿昌语、拉祜语、基诺语的影响。

4. 从受多少语言影响和时间上去考察

语言间的影响，有的可能受多种语言影响，有的可能只受一种语言影响，这跟语言接触面有关。像现代汉语就可能受多种语言影响，如从借词来看，就有英语、俄语、日语等的影响，从整个汉语来看，则受到西域少数民族语言、蒙古语、满语、印度语等的影响。颜之推在《颜氏家训·音辞篇》就描述了当时汉语受影响的情况，即所谓“北杂夷虏，南染吴越，皆有深弊，不可具论”，《颜氏家训》中也多次说到鲜卑语对汉语的影响。再像景颇语也受到汉语、傣语、缅语等多种语言的影响；像汉语河州话（分布于以甘肃临夏为中心的甘肃、青海部分地区）不仅受到藏缅语系语言的影响，还受到阿尔泰语系语言的影响。而像凉山彝语、广西壮语只受到汉语的影响。

一种语言受到别种语言影响的时间，有的容易说明，如有的语言的文字记载可以为证，如汉语至少从汉代起就受到别的语言的影响，汉语中所积淀的别种语言影响的痕迹较为复杂。有的不容易说明，如苗瑶语、壮侗语中的一些词，到底是汉语的借词，还是本族语固有的词，还是跟汉语同源的词，很难划清。

不过，被影响的语言中所保留的别种语言的痕迹及其历史层次，往往反映了民族关系和民族文化交流历史，包含了重要的历史文化价值，语言学、民族学、人类学、历史学、文化学等学科都非常重视这方面的研究。

语言影响的因素是多方面的，但主要有两个方面，一是社会因素，包括政治经济力量、文化特征、人口多少、民族心理、民族关系、宗教信仰等。一般来说，两个民族接触时，政治经济力量强大、文化先进、人口多的民族的语言，总是较多的影响另一个民族，如汉语在历史上总是较多的影响少数民族的语言，1949 年以后更是明显。汉语在历史上对日语、朝鲜语、越南语的影响，也是因为这个原因。另外，民族关系和谐、融洽时，语言影响的速度就快，否则，就会慢一些。宗教信仰也对语言接触有影响，如维吾尔族、哈萨克族因为信仰伊斯兰教，因而在维吾尔语和哈萨克语中就有阿拉伯语的借词。

与语言影响的深度和方式有关的另一个因素是语言文字本身，包括语言渗透能力的强弱，有无亲属关系，有无文字等。一般来说，渗透力强的语

言容易接受其他语言的影响；有亲属关系或类型相近的语言相互间容易发生影响，如在我国汉语对汉藏语系诸语言的影响就比对非汉藏语系的语言影响大些，同语系或同类型的语言之间，影响的成分容易被接受和吸收。没有文字的语言相对来说比较容易受有文字的语言的影响，如汉语在历史上对日语、朝鲜语、越南语的影响。

## 二、语言转用

语言转用(language shift)是指一个民族或一个民族的一部分人放弃使用母语而改用另一个民族的语言的现象①。语言转用是语言影响的进一步结果，它是随不同民族的接触进而融合而产生的。语言转用在中国历史上是较为多见的，春秋战国时期的所谓东夷、南蛮、北狄、西戎诸民族经过会盟、战乱、兼并、散居等，从而发生民族的融合和语言的融合，这些民族的人都转用了汉语；两汉以后，北方的匈奴、鲜卑、羯、氐、羌，隋唐以后的契丹、女真等民族也在同汉族的交往中被融合，汉语成了这些民族的共同语言。语言转用现象不仅发生在中国，世界上许多民族都曾发生过语言转用的现象，如美洲、非洲、澳洲等许多民族或土著在历史发展进程中放弃了本民族语言而改用了英语、法语、西班牙语等欧洲语言。

语言转用可以从类型、社会历史条件、语言自身的条件、自愿转用和被迫转用等不同角度去观察。

### 1. 语言转用的类型

由于各民族的社会历史条件不同，语言转用出现过不同的类型。有的是整体转用，即一个民族全体转用另一个民族的语言，如中国的回族转用汉语就属于整体转用。不过，中国的回族转用汉语后，回族这个民族还保留着，汉语成为汉族人和回族人的共同交际工具，而有的民族转用汉语后，民族成分也随之消失。如南北朝时的鲜卑族在中国北方建立了北魏政权，汉族和其他民族被鲜卑族统治，但由于经济、文化以及人口的因素，鲜卑族不得不向汉族学习，主动被同化，如北魏孝文帝主动提倡学习汉语，禁止在朝廷上使用鲜卑语，规定在朝廷上不能说“北俗之语”(即鲜卑语)，否则免官罢职，要求讲“正音”，即讲汉语。后来，鲜卑人转用汉语，其民族成分也融入汉族。再如西夏人、契丹人在转用汉语后，其民族也被汉族同化了。另一种转

① 语言转用也叫语言替换、语言替代、语言融合、语言排挤等。

用是主体转用,即一个民族的主体或大部分人转用另一种语言。如满族现有400多万人口,除黑龙江省爱辉县和富裕县极少部分人能说满语外,大部分人已转用汉语。再如土家族,现有人口570多万,除了湖南湘西土家族苗族自治州一些偏僻山区约20万人会说土家语外,大部分人转用汉语或苗语。像畲族以及新疆境内的乌孜别克族、塔塔尔族也都发生了语言的主体转用现象。还有一种是局部转用,指一个民族的某个局部地区转用其他语言,如东北地区的蒙古族人转用汉语,云南通海地区的蒙古族转用彝语,湖南省桃源、常德的维吾尔族转用汉语。语言的局部转用还有一些有趣的现象,一个家庭有的成员使用本族语,有的成员却转用其他语言,多数是青年人转用其他语言,如云南元江老巴阿村的哈尼族的年轻人大多转用汉语。

**2. 语言转用的条件**

语言转用是由社会历史决定的,具体说有民族的融合、分布的变迁、民族关系的变化等原因。语言的转用首先是由于民族的融合,两个或多个民族融合为一个民族时,需要一种共同的交际工具,而民族融合时哪一个民族的语言被别的民族所转用是由社会历史条件决定的,政治、经济、文化、生产力先进的民族的语言往往成为被融合的民族学习和转用的语言。如我国魏晋以后,北方许多少数民族如匈奴、契丹、鲜卑、党项、女真等,在同汉族融合时,都转用了汉语,这些民族有的在历史上曾经统治过汉族,但由于当时汉族的政治、经济、文化、生产力水平都较统治民族先进,客观上促使这些民族学习汉文化和汉语,在民族融合的过程中语言也转用为汉语。

分布的变迁也促使语言的转用。如满族转用汉语,虽然原因是多方面的,但居住环境的改变应该是一个主要原因。顺治元年清军入关以后,满族人由大规模聚居,变成分散到汉族居住,加上朝廷提倡学习汉语,就使得大部分满族人逐渐放弃了满语而转用汉语。历史上,汉族周边的一些民族转用汉语,都是因为这些民族从聚居转变为散居在人数占优势的汉族人中的缘故。杂居是语言转用的客观条件。像云南通海地区的蒙古族是因为战争的原因定居在云南的,并与当地的彝族通婚,这就促使他们发生语言转用,转用定居地的彝语。湖南桃源、常德的维吾尔族也是由于战争而定居该地,由于人口少,加上同汉族、回族等民族的杂居,就转用汉语了。

民族之间的通婚也能促使语言的转用,如我国的仡佬族有40多万人,但只有很少的人能说仡佬语,大多数人都说汉语。仡佬族语言转用的原因主要有两个,一是散居在汉族中间,二是与汉族人通婚,通婚后一般使用汉

语，或者经过双语阶段逐渐过渡到汉语。语言的转用大多经历了从双语到转用这样的过程。

3. 被转用的语言应具有什么样的条件

不同民族之间发生语言的转用，某个民族转用哪一个民族的语言往往是由人口的多少、经济文化的水平、军事力量的强弱、居住情况及语言自身的特点等诸多因素所决定的。通常由人口多少和经济文化水平起主导作用，鲜卑、女真、契丹、党项、满族等民族在历史上曾经统治过汉族，但由于人口少，经济文化落后，加上跟汉族杂居，逐渐被汉族融合，语言上也转用汉语。反过来，一些汉族人杂居在少数民族中间，也可能转用少数民族语言，如云南路南县、镇雄县的一些汉族人跟彝族杂居或通婚，有的已经转用彝语。

4. 自愿转用和被迫转用

就许多民族在转用汉语的过程来看，有些民族是自觉放弃本民族语言而转用汉语的，有些民族却是因为经济、文化的需求被迫转用汉语的。前者可以叫自愿转用，后者叫被迫转用。隋唐以前，匈奴、鲜卑、羯、氐、羌等民族有的在取得政权以前已经同汉族杂居，这些民族的大多数人已经学会汉语，他们在建立政权后往往把汉语作为不同民族之间的交际工具。像鲜卑族在建立北魏之前虽然没有同汉族杂居，但在中原建立政权之后，由于经济和文化的需要，北魏政权鼓励甚至强令鲜卑族学习和使用汉语。而隋唐之后，契丹、女真等民族在中原建立政权之后，曾采取措施阻止族人学习和使用汉语，如金世宗曾令皇室成员、卫士及其他女真人“不得学汉语”，不能用汉名，甚至设法不让女真人跟汉人杂居。但最后还是发生了民族的融合和语言的转用。满族在入关之后也曾训诫本族人不要忘记满语，但最后的结果是满族人在两三百年中大部分转用汉语。

## 三、语言兼用——双语现象

1. 双语的性质和分类

双语(bilingualism)或多语(multilingualism)是指某一语言社团(比如民族或国家)或个人同时使用两种或两种以上的语言。从这个意义上讲，双语包括了多语，下文介绍不再区分双语和多语。双语也就是语言接触中的语言兼用现象。

社会语言学家对双语进行了进一步的分类。社会中双语并用的情况叫社会双语，社会双语现象在许多国家或民族是普遍现象，有的国家，如非洲

的喀麦隆有几百种语言，许多人会说多种语言，马耳他人除了说马耳他语外，多数人还会说意大利语、英语及阿拉伯语。实际上，现代社会中要找一个纯粹单语的社会或国家是比较难的。

个人使用两种以上语言的情况叫个人双语。如我国云南西双版纳的基诺族都会说母语基诺语，但大部分基诺人也会说汉语。再如菲律宾人、马来西亚人除了说本族语外还兼说英语。个人双语是社会双语的具体表现。

国家的官方语言有两种或多种的叫国家双语。马耳他的官方语言是马耳他语，但个人或社会是双语或多语，这是国家单语、个人或社会双语。瑞士国家规定德语、法语、意大利语、罗曼斯语都是官方语言，但大多数居民只能说其中的一种语言，这是国家多语、个人单语。在菲律宾，菲律宾语和英语都是官方语言，居民也大多既会说菲律宾语，也会说英语，这是国家双语，也是个人社会双语。

2. 双语现象种种

双语现象是语言接触的结果，但双语现象是一个复杂的语言社会现象，考察双语现象的各种复杂情况有利于认清双语现象的规律和语言接触的具体情况。

在语言接触中发生双语现象，有的是全民性的，有的只是局部性的。全民性双语是指一个民族或社会的全部和大多数人都操双语，如上文说到的基诺族都会说自己民族的母语，但也广泛使用汉语，他们用汉语跟别的民族交往，听汉语广播，看汉语电视电影，阅读汉文书刊，在基诺人那里，汉语跟基诺语一样，都是他们生活、工作和学习的必要的交际工具。其他像鄂温克族、达斡尔族、裕固族、京族的大多数人都能使用本族语和汉语。局部性双语是指一个民族的一部分操用双语，如聚居在大小凉山的彝族人一般只会彝语，但散居在云南、贵州、广西的彝族人，由于跟汉族人杂居，多数是操彝语和汉语的双语人。再如贵州各地的瑶族人除了说瑶语外，一般成年人还会说汉语和布依语。

操双语的人有的能熟练使用双语，有的则不能熟练使用双语。我国一些少数民族兼用汉语时，只能对付日常生活用语，不能进行较为深入和复杂的交际。从兼用何种语言来看，多数是少数民族兼用主体民族的语言，如我国许多少数民族中都有兼用汉语的情况；也有一个少数民族兼用另一个少数民族语言的情况，如云南怒江州中杂居在傈僳族中的其他少数民族都有兼用傈僳语的现象，这是因为傈僳族在该地人口多，分布广，语言较为一致。

再如，云南贡山县独龙族、藏族、怒族、傈僳族、纳西族杂居的地方，各民族之间常常有多语兼用的情况。主体民族兼用少数民族语言的情况较为少见，在我国少数民族地区，有些汉族人杂居在其中，这些汉族人往往学会使用当地少数民族的语言，如四川阿坝藏族自治州的理县汉族人兼通羌语，云南芒市的不少汉族人会说傣语。

3. 双语形成的原因和双语能力

双语是语言接触的结果，民族杂居、民族接触、移民、文化交流和教育目标、政治经济因素等都可能形成双语。

民族杂居最容易形成双语，因为不同民族杂居在一起生活，必然发生不同民族之间的交往，需要学习对方的语言，客观上也提供了互相学习对方语言的条件。我国各少数民族中凡是跟别的民族杂居的，双语现象就比较普遍，双语程度都比较高。移民也会引起双语现象的发生，一个人口相对较少的民族迁徙到人口较多的民族中，很容易形成双语，如东南亚的一些苗族人近几十年移入美国，青年人几乎都成了会说苗语和英语的双语人。在一个多民族国家，民族交往和接触比较普遍，一些人口较少的少数民族因为经济、政治的原因往往愿意学习人口较多的经济文化比较先进的民族的语言，形成一种普遍的双语现象。我国许多少数民族都会说汉语，就是因为少数民族是“小聚居”，跟汉族人是“大杂居”，汉语成为少数民族兼用的第二语言，也是各民族之间共同使用的交际工具。

现代社会中，双语的形成跟双语教育有一定的关系。1949 年以后，为了尽快发展我国少数民族地区的经济和文化，为了使少数民族学生能受到高层次的教育，我国在少数民族地区实行了学生自愿自择前提下的双语教育，即鼓励、提倡学习本族的语言文字，也积极开展汉语文教育。双语教育使得少数民族中的干部、知识分子、青年学生大多数不同程度地学会了汉语。当前随着世界经济全球化和一体化的到来，随着中国加入世界贸易组织，越来越多的中国人重视了对英语的学习，从幼儿园、小学、中学到大学，有条件的都在提倡双语双文化教学，也许不久的将来，在中国操汉英双语的人会越来越多。这种双语现象的形成是双语教育的结果。

双语能力往往跟年龄、性别、职业有关。据统计在全国操双语的各少数民族中，7～45 岁的双语人明显多于 46 岁以上的人，如苗族 7～17 岁的双语人占 7.06%，18～45 岁占 12.63%，46 岁以上仅占 5.08%；壮族 7～17 岁的双语人占 15.1%，18～45 岁占 29.48%，46 岁以上仅占 10.14%；布依

族 7～17 岁的双语人占 16.05%，18～45 岁占 26.86%，46 岁以上仅占 11.56%；黎族 7～17 岁的双语人占 15.20%，18～45 岁占 24.63%，46 岁以上仅占 9.45%；达斡尔族 7～17 岁的双语人占 22.17%，18～45 岁占 34.26%，46 岁以上仅占 10.46%。而维吾尔族、哈萨克族、独龙族、基诺族等民族 60 岁以上的人几乎没有双语人。可见，双语人的年龄规律具有“前高后低”的特点①。就双语人的性别来看，男性高于女性，据统计全国少数民族双语人的性别比是 134.47（性别比＝男性人数/女性人数×100）②。双语能力跟职业和受教育的程度也有关系，中国少数民族双语人中，干部、知识分子、学生的双语能力要高于一般群众。

**4. 双语教育**

双语是社会现象，是民族接触和语言接触的结果，同时，在当代社会中，双语也是一个教育问题。在中国，双语教育分两个层面，一是少数民族的双语教育问题，另一是汉语和外语的双语教育问题。关于中国少数民族的双语教育，应该注意以下几个问题。

一是要处理好少数民族语文教学和汉语文教学的关系。凡是通行少数民族语文的地区，原则上应该在掌握本民族语文的基础上再安排汉语文的学习。儿童入学后，先以母语作为教学语言，有本民族文字的，课本用本民族文字编写，同时开设汉语课。本民族没有文字的，用汉语汉字编写教材，并从入学时就开始汉语文教学，但要处理好如何用本民族语言讲课的问题，等到儿童掌握了一定数量的汉字，汉语达到一定的水平时，再转为以汉语作为主要教学语言。

二是在汉语教学中，不能照搬对母语为汉语的学生所使用的方法，要注意少数民族学习汉语的特点，尤其要针对他们母语的不同特点来进行教学。这方面可以借鉴对外汉语教学的经验。中国境内的少数民族使用 80 多种语言，这些语言分属 5 个不同的语系；不同语系之间，同一语系的不同语言之间，各有不同的语言特点，母语特点会对汉语学习形成干扰。只有针对学生母语的不同的特点进行针对性的汉语教学，才能提高汉语教学的质量。

三是要重视对教学对象的研究。不同的少数民族的实际情况不同，有的民族如藏族、维吾尔族、哈萨克族人口相对较多，有大片聚居地，懂汉语的人少；有的民族如白族、纳西族、阿昌族人口相对较少，跟汉族人杂居，懂汉

①② 周庆生.《中国双语人口构成》.载《双语双方言》(五).汉学出版社，1997 年

语的人多；有的民族如蒙古族、藏族、维吾尔族、朝鲜族、彝族历史上就有自己的文字；有的民族如独龙族、怒族、德昂族、布朗族没有自己的文字；有的民族如景颇族、拉祜族、傈僳族只有新创制的文字。这些民族在学习汉语时的情况各不相同，应该根据不同的教学对象，确定教学目标和教学方法。

四是要培养出一大批既熟悉汉语，又熟悉少数民族语言，既懂得双语教学的特点，又了解国家民族政策和民族语文政策的双语教学教师。

关于汉语和外语的双语教育，也要注意几个问题。

一是要在学习好汉语文的基础上进行双语教学。目前许多大中城市都开办有不少双语幼儿园、双语小学或中学，甚至有的学校用英语教语文（中文），这一方面说明，社会、学校、家长十分重视外语学习，但在幼儿阶段过分重视外语学习，势必会影响汉语文学习，影响孩子的汉语文听说读写能力的培养。双语教学要处理好母语汉语文学习跟外语学习的关系，不可偏废。

二是要重视双语教学的平衡性。目前我国的外语教学并不平衡，东部地区好于西部地区，大中城市好于农村或偏僻地区。目前应加强西部地区、农村和贫困落后地区的外语教学。

三是要培养出懂得双语教学规律和教学法的双语教师。目前在部分声称进行双语双文化培养和教学的幼儿园、小学、中学，实际上多数是增加了英语课程，并不是真正的双语教学。原因是这些幼儿园和学校没有专业的双语教师。

## 四、语言混合——皮钦语和克里奥尔语

两种语言相互接触时还可能发生语言的混合。语言的混合是语言接触所产生的特殊形式，语言混合主要有两种形式——皮钦语和克里奥尔语。

### 1. 皮钦语

皮钦语（Pidgin）也叫洋泾浜语。洋泾浜原是上海外滩的一段，位于叫做洋泾浜的河流（早已填没）和黄浦江的汇合处。1840 年以后，洋泾浜一带成为外国商人聚集的地方，当地人（据说尤其是人力车夫）跟外国人就用一种支离破碎的英语通话，如 comprador（买办）说成“康白渡”，last car（末班车）说成“拉司卡”，all same（全部）说成“何洛山姆”，one dollar（一块大洋）说成“温淘萝”，把“好久不见”说成“long time nosee”。后来人们就把“洋泾浜”作为这种支离破碎的外语的中国名称。洋泾浜的学名叫皮钦语（Pidgin），是中国人说洋泾浜英语“business”的讹音。

皮钦语不仅中国有,它是世界上许多通商口岸的一种常见的语言混合现象,是通商口岸的当地人在跟外来的商人、水手、传教士等打交道过程中学来的一种变了形的外语。中国的洋泾浜英语早在18世纪中期就有记载;20世纪70年代、80年代,在德国的柏林和法兰克福,由于大量的土耳其、意大利、西班牙和葡萄牙等客籍工人的流入,产生了以德语为基础的皮钦德语;在巴布亚新几内亚有皮钦英语。皮钦语的共同特点是语音经过当地语言音系的适当改造,语法规则减少到最低限度,词汇项目比较少,交际中往往要借助于迂回曲折的说法指称事物。

皮钦语是一定社会条件下的语言混合物,能说皮钦语的人都有自己的母语,皮钦语只是在跟外国人打交道时的特殊场合下使用的一种变形外语。由于皮钦语不是母语或第一语言,交际的特殊场合没有了,皮钦语也就会消失。像中国上海的洋泾浜在20世纪50年代以后,失去了社会基础,也就不再被使用了。皮钦语从发展趋势来看,主要有两个趋向,一是随着其存在的社会基础的消失而消失,二是发展成为另外一种混合语——克里奥尔语。

2. 克里奥尔语

皮钦语不是母语,也不是人们的第一语言,没有人把它当作母语去学习。如果皮钦语成为第二代人学习的母语,成为社会采用的主要交际工具,就产生了另外一种混合语,即克里奥尔语(Creole)。如非洲某些地区的种植园,欧洲殖民者和非洲劳工之间、劳工和劳工之间不能通话(劳工来自不同的部落),惟一通用的交际工具只能是经过洋泾浜化的殖民者的语言,这种混合语被不同种族、不同部落的人所使用,并被作为母语被下一代学习,这样洋泾浜语就提升为克里奥尔语。另外,像海地有以法语为基础的克里奥尔语,牙买加有以英语为基础的克里奥尔语。据估计世界上只说克里奥尔语的人多达1 700多万人。

与皮钦语的发展方向不同的是,克里奥尔语一旦形成,其词法、句法结构比较丰富,语音也开始规范,词汇量和语言的功能也变得丰富,整个语言体系变得越来越稳定,最后变得跟其他语言同样完备,是一种完整的语言。

从社会语言学角度看,克里奥尔语是语言接触的特殊形式,克里奥尔语并不是一种低劣的语言。有学者认为,现在世界上任何一种语言都可能在历史上某个时期曾经是一种克里奥尔语,特别是一些使用人口较多的语言,其历史上总有许多语言接触导致混合的现象。有人甚至认为现代英语与其前身盎格鲁撒克逊方言结构系统上有很大的不同,这是英语在历史上跟丹

麦语、法语及其他许多语言接触的结果，因而现代英语是从一种克里奥尔语发展而来的。这种说法是不无道理的。

不过，克里奥尔语总是被人认为是某个标准语的坏的变种，因而随着讲克里奥尔语的社团的发展和分化，社团的一部分人就产生了向相关的标准语靠拢的愿望和社会压力，而随着社团成员的不断努力，这种克里奥尔语也就可能变得越来越像相关的标准语，这样就会产生所谓的“克里奥尔的脱化”(decreolization)，如现在的牙买加英语就是一种脱化的克里奥尔语。通常在具有脱化的克里奥尔语的语言社团中，会出现层次化现象，即社会上层使用的语言最接近相关的标准语，社会阶层越低的人保留原来的克里奥尔语的特征越多。由此也可以看出，社会阶层对语言变化的影响。

### 五、语言冲突和语言地位规划

在一个多民族国家或社会中，不同的民族使用不同的语言。由于语言跟民族有着不可分割的关系，语言是民族的标志，许多民族都非常重视保护、推广本民族的语言，这样在一个多语言国家或社会中必然会发生语言冲突，尤其在语言地位规划中，这种语言冲突就会表现得十分强烈。语言地位引起的语言冲突实际上是民族地位和社团地位冲突在语言上的反映，每一个民族或语言社团都想为自己的语言争取一定的语言地位。

说起语言冲突，最为典型的恐怕就是加拿大的法语和英语之争了。加拿大是典型的移民国家，居民来自不同的国家和地区，其中主要来自英国和法国，英裔占 40%，法裔占 27%，从其他国家和地区来的人占 31%，土著只占 2% ，全国多数人讲英语，少数人讲法语。加拿大宪法规定，英语和法语都是官方语言，所以加拿大属于国家双语国，国家或政府、议会的法律、法令、通知、文件同时使用英语和法语。但其中的魁北克省是讲法语的人集中的地方，80%的人讲法语，因此，魁北克省的法语意识特别强烈。该省的法语宪章规定，法语是该省的工作、教学、交通、邮电、商业和科学用语。法语由于受到英语和英语文化的包围，致使一些人的法语保护意识特别强烈，甚至出现过激措施来保护法语，如 1977 年魁北克省通过了著名的《101 号法律》，规定法语为该省议会惟一的正式语言，法语文本为惟一的正式文本，一切商业广告只能使用法语。这个法律引起魁北克省讲英语的人强烈不满，于是省议会又通过了折中的《198 号法律》，规定室外商业广告只能用法语，室内广告可以使用其他语言。这样，这场语言冲突或矛盾才缓和下来。魁

北克省的语言冲突还引起了政治运动，随着魁北克省法语势力的扩张，该省的独立运动也日益发展起来，以致1995年举行了全省公民投票，公决是否脱离加拿大(因未达到规定票数而未果)。

白俄罗斯曾是前苏联主要加盟共和国之一，1990年之前官方语言是俄语和白俄罗斯语。前苏联解体之后，1990年白俄罗斯政府实行新的语言政策，大力推广白俄罗斯语，规定政府和机关的公文只准用白俄罗斯文印发，国家电台、电视台只准用白俄罗斯语播音，各种票据只用白俄罗斯文印刷，并于1994年在白俄罗斯宪法中规定白俄罗斯语是白俄罗斯的国语，国家充分保障公民把俄语作为民族间交际语的权利。这样就使得俄语失去了官方语言的地位，下降为民族间交际语。政府还强制使用白俄罗斯语作为教学语言，报纸中增加白俄罗斯语版面。白俄罗斯如此推广白俄罗斯语，当然是从政治和民族的独立性角度考虑的，但当时的实际情况是，在日常交际中大多数白俄罗斯人仍然使用俄语。这样，俄语是否成为白俄罗斯的国语，引起了广泛争论，双方争执不下，于是于1995年举行全民公决，结果是81.3%的选民支持把俄语作为国语，与白俄罗斯语具有同等的官方地位。到此，进行了5年之久的“推广白俄罗斯语运动”和“国语之争”才平息下来。

二战以后，亚洲许多民族从殖民地手中独立，独立之前，殖民者的语言往往具有官方语言或实际的官方语言的地位，殖民者语言尤其在教育、行政、法律中有权威地位。民族独立后，独立的国家往往进行民族语言地位的独立，力争使民族语言成为官方语言。

1947年印度独立，1950年成立印度共和国，印度共和国是一个典型的多民族、多语言、多文字的国家，但英语一直是印度的行政和教育语言。印度独立后非常重视民族语言的地位规划，印度宪法规定有14种法定语言，其中天城体字母书写的印地语是惟一的联邦官方语言，另外梵文、乌尔都语也是不限定地区的法定语言，另外的11种语言都是各邦的官方语言。1950年政府规定15年之后，印地语作为官方语言正式实施，同时使用英语，但到1963年(离规定只有2年)，由于印地语作为官方语言的条件还不成熟，国会只好决定延长英语的使用期限，一直使用到不需要英语为止。不过，把印地语作为母语和第二语言的印度人占全国人口的大多数，印地语成为印度全国的共同语还是很有希望的。

非律宾历史上曾采用过阿拉伯语、西班牙语、英语作为官方语言或通用

语言。1935 年菲律宾成立自治政府,1946 年菲律宾独立。独立后,菲律宾决定采用以他加禄语为基础的菲律宾语为国语,20 世纪 70 年代之后进一步规定他加禄语为官方语言和国语。但由于英语在菲律宾被广泛运用于各种领域,所以,从 1974 年开始,菲律宾教育部不得不推行菲律宾语和英语双语制。双语制在菲律宾是合乎实际的,因为菲律宾语在菲律宾的行政、教育、科技等领域的普及率还不够高。

新加坡的语言问题更加复杂和敏感,新加坡也是一个多语言的地区,讲英语、汉语、马来语和泰米尔语。1965 年新加坡共和国独立,宪法规定马来语、华语(汉语普通话)、泰米尔语、英语 4 种语言为官方语言,其中马来语是国语,华语是华人的民族共同语,泰米尔语是新加坡印度人的民族共同语,英语是新加坡行政、教育、贸易和各族人民间的交际语言。新加坡在语言冲突问题的解决上走以英语为主,多语并存的双语制(多语制)的语言政策,对语言冲突的解决是很有启发的,它一方面实行双语制(多语制),一方面也不忽视作为第二语言的民族语言,同时又防止了民族语言运动盲目地发展而成为民族沙文主义。

在一个多民族多语言的国家内总有语言冲突,因而多民族、多语言国家都有语言地位规划的重要任务,因为语言平等是民族平等的一个重要方面。我们认为在多民族、多语言国家解决语言冲突,进行语言地位规划时,要从实际出发,一方面要实现语言平等以体现民族平等,另一方面要重视各民族的共同交际语地位规划,以有利于各民族的团结和沟通。这方面,印度、新加坡、菲律宾、以色列等国的语言地位规划就是很好的例子,他们一方面重视民族语文的地位,另一方面重视英语作为各民族的共同交际语的地位,英语地位的保留,也有利于跟国际社会的交往。那种所谓"一个国家,一个民族,一种语言"的民族意识是十分狭隘的。

我国的语言政策也是十分成功的。我国是一个以汉族人为主的多民族国家,语言上采取的是以汉语普通话为全民共同语,各民族语言平等发展的语言政策,《中华人民共和国宪法》既明确规定"国家推广全国通用的普通话",也明确规定"各民族都有使用和发展自己的语言文字的自由"。《中华人民共和国国家通用语言文字法》既明确规定"国家推广普通话",普通话是国家通用语言,也明确规定"各民族都有使用和发展自己的语言文字的自由"。在《中华人民共和国民族区域自治法》中也规定:"汉族干部要学习当地少数民族的语言文字,少数民族干部在学习、使用本民族语言文字的同

时，也要学习全国通用的普通话和汉文。”

## 第三节 微观社会语言学

### 一、语言变异和语言变体

微观社会语言学的经常性课题是语言与社会、语言与阶级、语言与环境、语言与性别、语言与种族集团等，实际上它是讨论各种社会因素对语言变异的影响。

语言变异(language variation)是指语言因各种社会因素影响而产生的种种差异，它既包括某个特定语言项目在共时分布上的差异，也包括该语言项目在历时流程中的变化。变异理论是微观社会语言学的基本观点，社会语言学家把语言放到社会环境中去研究，具体分析语言的各种变异，认为语言不是一种同质系统，而是一种有序异质结构，语言研究者通过变异成分的分析来考察变异同社会因素之间的共变关系。

跟语言变异相关的一个概念是语言变体。语言变体(variety in language)是社会语言学研究的对象单位，是指具有共同的社会分布的一组语言项目(语言形式)，语言项目指的是各种语言成分或规则，可以指整个语音系统、语法系统、词汇，也可以指一个词、一个音位或音位变体，社会分布则指相应的语言社团，一群接受和使用某一些语言项目的人。社会语言学用“变体”这个概念来描写语言的变异性特点。

语言的变异，有的跟地域有关，有的跟说话人的社会身份(包括年龄、性别、社会地位、职业等)有关，有的跟使用语言的场合有关。因而，语言的变体可以分为语言的地域变体、语言的社会变体、语言的功能变体①。每一种变体下面还可以分为不同的小类。

语言的功能变体即一般所谓的“语体”，是以语言交际功能为标准所划分的语言风格类型，是适应不同交际目的、内容、范围的需要在特定语境中运用全民语言而形成的语言功能变体。常见的功能变体有口头语体和书面语体，书面语体再分为政论语体、科技语体、公文语体、文艺语体等。关于

① 也有二分的，分为社会变体和功能变体，因为地域变体也是说话人的社会特征之一。

"语体"一般现代汉语教材、语言学概论及修辞论著中多有涉及,本章不再赘述。本章将重点介绍语言的地域变体和社会变体。

## 二、语言的地域变体

### 1. 地域变体与社会

语言的地域变体当然跟地理因素有关,由于地理上的距离才会造成不同的方言,由于地理上联系,才会形成语言联盟。但地理因素只是造成方言或语言联盟的条件之一,有时甚至不是主要的或起决定性作用的条件。方言或语言联盟的形成实际上也是由于社会原因产生的。比如四川省和重庆市基本上属于北方方言的西南次方言区,其入声一般都消失了,可是离重庆市区不远的江津县(现划归重庆市)还有入声,这是因为江津地处偏僻,受官话影响小,与成都、重庆接触相对较少,语言发展缓慢,语音中就保留了一些中古汉语的特征。再如湖南邵阳地区各个县的方言由于受到西南方言的影响,入声基本消失,但邵阳市区还保留了入声,这又是为什么呢?从地理上也不好解释。原因是邵阳是湖南省的一个历史悠久的文化古城,在湖南的经济文化上都占有重要的地位,这样邵阳人虽然较周围各县的人有较多的机会接触别的方言,但邵阳在经济和文化上的权威地位使得邵阳在抵制别的方言的影响上具有更多的排斥力。再如同为闽语的莆仙话和闽南话虽然地理分布上是紧邻的,但彼此不能通话。这种现象也跟社会历史有关,比如说北方人南下的时间先后不同,先后带来的北方话自然有差别,或者是南下前北方话内部或与北方话接触之前的当地古越语内部本来就有差异,也可能是南方各方言内部独立发展出的一些新成分。

几种不同的语言长期共存于一个地区有可能产生语言联盟,但像瑞士,尽管几种民族语言长期共存,可是一直没有形成语言联盟。可见,地理因素在形成语言联盟方面也并不都起决定性作用。

### 2. 方言

传统上对语言地域变体的研究也就是方言学的研究。方言(dialect)广义上的理解包括地域方言(regional dialect)和社会方言(social dialect),社会方言也就是语言的社会变体。狭义的理解,方言专指地域方言。

地域方言是语言在不同地区的地域变体。一般来说,同一个地域方言分布在同一个地区或同一个地点。同一种地域方言也有在地理上不相连接的,从而形成方言岛(speech island)或语言飞地(outlier),这主要是由移民

造成的。

方言是语言在地域上分化所得出的变体，一种语言可以分化为若干种方言，方言还可以不断分化，形成不同的层次：语言→方言（区）→次方言（subdialect，片）→土语（vernacular，小片）→腔（accent，点）。如闽方言是汉语的一种方言，可以分为闽北、闽南、闽东、闽中、莆仙、琼雷 6 个次方言①，闽南次方言又可以分潮汕土语、泉漳土语、大田土语，泉漳土语又可以分出泉州腔、漳州腔②。再如吴语区是汉语的一种方言，下面可以分出几个片（宣州片、太湖片、台州片、瓯江片、婺州片、处衢片），每片可以分出多个小片，每个小片可以分出不同的方言点。

同一种语言分化到什么程度才算是不同的方言？同一种方言分化到什么程度才算不同的次方言？这不仅仅是由方言自身决定的，而且与政治、民族、文化、地理等因素有关。方言形成的原因是复杂的。其中主要是社会分化造成的，如我国的客家方言就是由于战乱等原因，中原汉人（即所谓客籍，如山西上党，河南弘农、安丰和新蔡，安徽寿县）数次南迁，引起汉族社会的分化而产生的。如果把美国英语、澳大利亚英语看作欧洲英语的方言，其演变也是由于社会分化形成的。大规模的移民跟所在地的语言或方言互相接触、融合也是形成方言的重要原因。如果从语言自身的差别（是否听得懂）来看，汉语方言之间的差别要大于法语和西班牙之间的差别，也大于荷兰语和德语之间的差别。靠语言自身的差别来判断是否是不同方言，是否是不同语言，往往是不可靠的。粤语同北方话的差别很大，但还是同一种语言的不同方言。国外一些语言学家认为中国的方言是不同的语言，这是不正确的，因为除了语言自身的标准外，最重要的标准是社会政治标准，而且汉语各方言有共同的书面语。

不过，同一种语言到底有多少种方言往往是有争议的，如汉语方言就有 7 种说、8 种说、10 种说的不同。一般来说，北方方言（北方话、官话）、湘方言、赣方言、吴方言、客家方言、粤方言 6 种方言是没有争议的，有人将闽方言分作闽南和闽北两个方言区，这就是八大方言区，不分则是七大方言区，《中国语言地图集》中多了徽语、晋语、平话 3 种，这就是十大方言区，也有分为 3 个大区：北部方言（官话）、中部方言（吴语、湘语、赣语）、南部方言（闽

① 参阅侯精一主编.《现代汉语方言概论》.上海教育出版社，2002 年。

② 参阅游汝杰.《汉语方言学导论》.上海教育出版社，2000 年。

语、粤语、客话)。目前多数论著,尤其是一般教材,都采用七大方言说。

3. 现代汉语方言的形成

现代汉语七大方言的形成都有深刻的社会历史原因。北方方言可以粗略地看成是远古中原一带华夏、东夷等族的语言长期自身发展,相互影响,跟四周方言交融,并向四周延伸的结果。我国历史悠久,是人类文明发源地之一,几十万年以前,黄河流域就有先民居住,是中华民族文化产生的源头。从甘肃、青海向东,尤其是陕中、晋南直到冀鲁豫等中原官话和冀鲁官话等的分布地区,自然条件优越,黄河等河流流入该地,使这里土地肥沃,再加上气候温和,宜于农畜,于是周边地区不同民族的人民都向这里移动,使这个地区成为我国原始文化孕育和发展的中心地区,也是汉语从形成到发展的中心地区。不过,活动在这个中心区的民族主要是华夏民族,夏部落是由10多个大小近亲氏族部落发展而来的,和夏部落结为联盟的还有它的远亲氏族部落,以及东夷的一些氏族部落。事实上,东夷各部落很早就居住在齐鲁大地,跟华夏民族长期交流和融合,可以说很早就成为华夏文化的一部分。华夏族是后来汉族的主体,他们使用的语言也应是汉语的前身①。

汉语方言的分歧由来已久,但频繁的战争、诸侯国的分合变迁、诸子学说的传播,也为方言的融合提供了条件,历经夏商周三代,华夏族活动的地区的语言逐渐成为汉语共同语的基础,形成所谓的"夏言"、"雅言"。后经两汉、三国、魏晋南北朝的战乱和大规模的人口流动,北方汉语进一步融合和交融,经隋唐宋三代的稳定发展,内部一致性大为增强,于是北方话作为一个方言大区在唐宋时代逐步明确起来。

北方话随着历史上的迁都、移民、屯田、戍边等社会历史的变化,逐渐向四周扩散,从而形成不同的次方言区,如华北东北次方言、西北次方言、西南次方言、江淮次方言等②,其中不同原因的大规模移民是各个次方言区形成的直接原因。这些移民跟当地不同民族或当地人语言亲密接触,就形成了不同特点的北方话的次方言区。

其余六大方言的形成,概括地说也是北方汉人不断南迁而逐步形成的,北方汉人在不同时期迁往南方不同地区,北方话就与当地语言融合形成了

① 郭沫若.《中国史稿》.人民出版社,1962年。

② 有的论著对北方方言次方言划分较细,如北京官话、东北官话、胶辽官话、冀鲁官话、中原官话、兰银官话、西南官话、江淮官话等。另外,不同次方言区的形成各有不同的具体原因。

南方相互歧异的六大方言。如吴语是在古代吴越方言的基础上融合北方汉语形成的，古吴越两国都是中原华夏民族的后裔（至少王族来自中原），跟江南土著百越部族生活在一起。古吴越方言是当时华夏语言的分支，跟百越族的“越语”不同。春秋时代，吴国和越国长期战争，古吴语和越语逐渐融合，形成古吴语。东晋及南朝以及再后来的南宋时代，由于战乱，大批北方汉人南下，使古吴语受到中原汉语的强烈冲击和影响，给吴语的发展注入了新的因素。现代吴语在宋元之际基本形成。

湘语是在古代湘语的基础上逐步形成的，古代湘语即春秋战国时所谓“楚语”。现代湘语的形成主要受到历代移民的影响，五代以前的外地移民主要来自北方，五代以后的移民主要来自东方，其中以江西为主。赣方言区在春秋时代处于吴、越、楚的交界处，东晋末年中原汉人南迁至江西中部，后来客家人多次南迁也以江西为中间站，从而逐渐形成赣方言。赣方言很受江淮方言和湘方言的影响。粤方言和闽方言的所在地区古代也属于百越，秦汉以后北方汉人南迁，并受当地语言的影响从而逐渐形成现代粤方言和闽方言。由于南迁的时间和迁入地的不同，形成了不同的次方言。客家方言的形成更是典型的移民所致。

可见，现代汉语方言及其次方言的形成除了地理因素外，更多的跟社会历史有关，尤其跟移民有关。大规模的移民跟所在地的方言或民族语言相互接触、融合，就逐渐形成新的方言。一部方言形成史也就是一部移民史，同时也是一部不同民族、不同语言或方言接触、融合的历史。

4. 共同语、标准语与方言

方言分歧的民族或国家、社会群体还会有共同语。如上古时期，存在汉语方言的分歧，所谓“五方之民，言语不通”，周秦时代政府就开始调查方言，汉代扬雄的《方言》就是确证。但当时也有所谓夏言、雅言、通语、凡语，夏言就是西周王朝都城镐京（今西安市附近）一带的方言，这种方言就是共同语的基础方言，这就是说当时华夏族活动的地区的语言是汉语共同语的基础。前文说过，北方方言在隋唐时期基本形成。明代明成祖定都北京后，大量的移民从山西、山东等地区移入北京，北京话接受了中原汉语的影响，加上北京逐渐成为全国的政治、经济、文化的中心，以北京话为代表的北方话逐渐成为汉民族共同语的基础方言。

共同语多数是基于本族语的某个方言之上的，如汉民族共同语基于北方方言，英语共同语的标准形式是基于伦敦方言，俄语共同语基于莫斯科方

言。也有一些民族共同语是外族语，如南美的巴拉圭的国语是瓜拉尼语和西班牙语，后者对土著居民来说是一种外族语。再如，我国的裕固族的民族共同语是汉语。

共同语跟标准语不同，共同语只是作为使用不同语言变体的人们相互交往的交际工具，并不要求使用者平常的口语以它为规范。标准语不仅是各方言地区之间相互交际的共同工具(这一点跟共同语一致)，而且是各方言的标准，具有较高的社会威望，是可用于一切正式场合的有书面语形式的语言变体(这一点不同于共同语)。标准语往往是人为地干预语言的结果，标准语的形成就是某个方言逐步标准化的过程。如普通话作为汉民族共同语实际上就是标准语，汉民族共同语是在北方方言的基础上形成的，也是逐步标准化的。标准化的过程包括对其语音、词汇、语法系统进行规范化，也包括扩大它在书面形式上的使用范围。汉民族共同语标准化经过白话文运动、国语运动，一直到 20 世纪 50 年代才确立其标准：以北京语音为标准音，以北方话为基础方言，以典范的现代白话文著作为语法规范的普通话。普通话无论是在口语和书面语上，还是在汉语各方言中都有着崇高的社会威望。

**5．语言联盟**

同一种语言因为方言分歧往往会产生共同语或标准语；几种不同的语言在同一个社会中共存，互相接触、互相影响，最后融合，一种语言往往会排挤、替代另一种语言；而有时几种不同的语言长期共存于某个地区或相邻的地区，互相影响，从而在语言结构上产生共同的特征，即几种不同的语言共有某些结构特征，不仅不同的语言继续并存，而且有时很难说哪一种语言影响了哪一种语言。这种语言现象就叫语言联盟。

语言联盟可以表现为某些语音、语法实体成分的共同，如非洲南部几种不同语系的不同语言都有一般语言所没有的利用吸气发出的吸气音(或称“喀嗒音”)。再如巴尔干半岛上的保加利亚语、马其顿语、罗马尼亚语、阿尔巴尼亚语属于不同的语族，但具有一些共有的结构特征：有后置词，动词不定式形式消失，属格和与格[①]合并，借助于意义为“想、要”的助动词构成动词将来时的分析形式。

① 属格即所有格、生格，主要表示领属、来源之类的语法意义。与格，主要表示动作的间接客体，或表示受某些介词支配的关系。

语言联盟的出现跟地域有关，不同语言共处一地或相邻，才可能出现语言联盟。语言联盟出现的内在原因，多数人认为跟大量的借词有关，如越南语跟汉语有联盟现象，就可能跟大量借词有关。不过，语言联盟出现的原因，现在研究和调查得还很不够。

## 三、语言的社会变体

语言的社会变体也称社会方言。语言的社会变体的研究是近几十年来才受到重视的，社会语言学家发现与语言变异最明显相关的社会因素是文化、性别、年龄、社会阶层、民族和种族。社会语言学的兴起就跟语言的社会变体研究有直接的关系，拉波夫对纽约市居民的发音的调查是语言社会变异研究的典型案例，拉波夫也开创了语言的社会变异研究。

### 1. 拉波夫的纽约市居民的发音调查

纽约人在说像 car、cart、flour、fourth、floor 等词汇时，有时候发出/r/的音，有时候则不发出这个音，至于什么时候发，什么时候不发，过去没有找到规律，认为这属于“自由变异”。拉波夫(1966)经过广泛收集语料，并结合讲话人的社会地位和讲话场合来进行定量分析，最后发现/r/变异看似杂乱无章，实际上背后有很强的规律性，这种规律概括起来讲就是，社会经济地位越高的阶层，讲话中卷舌音/r/的比率就越高；同时，无论是何种社会阶层，在越是正规的场合和注意自己语言给人的印象的场合，他们讲话时的发音中卷舌现象出现的比率就越高。可见，卷舌变异跟社会阶层和社会场合有关，拉波夫把他在纽约调查的讲话人归入了几个不同的社会经济等级，即“社会阶级”(social class)。拉波夫把纽约市民的社会阶层分为 6 个等级：上中等、下中等(以上为“中等阶级”)、上层工人、中层工人、下层工人(以上为“劳工阶级”)、低层(低层阶级)，卷舌变异跟社会阶层的相关性十分显著。这就是拉波夫对社会语言学的重大贡献，他根据对社会变异在城市社团的分布的调查结果，提出了一个崭新的语言社团概念，语言社团的一致性就表现在相同的语言态度上，而不是像过去所认为的那样，表现在语言社团成员的语言行为上。

### 2. 语言与文化

引起语言变异的社会因素很多，但语言跟文化的关系是最容易引起人们兴趣的一个问题。语言跟文化有着密切的关系，这是毋庸置疑的，但这个问题也最容易引起争议，比如是文化决定语言，还是文化影响语言；是语言

记录、记载了文化，还是语言影响了文化；文化是什么，语言是什么等等问题已经有不少争议。

(1) 语言与文化的关系

关于“文化”首先就有不同的理解。一般认为文化可以分为物质文化和精神文化两大类，而狭义的理解只指精神文化。房屋、器皿、机械、道路、工具等有形具体的东西属于物质文化，科学、技术、知识、社会组织、制度、政治、法律、伦理、道德、风俗、习惯、语言、教育、宗教、信仰、审美意识、文字、艺术等等都属于精神文化。

语言跟文化的关系，可以从不同的方面来理解。首先，可以认为语言是属于精神文化的，文化既然是人类以自身的能力在社会中创造出的成果，语言又是人类创造出来的交流思想感情的工具，语言自然应当属于文化。这是从两者的共同性来看的。

另一方面，语言同文化还有许多不同之处，如精神文化大多有阶级性，但语言没有阶级性，语言一视同仁地为各个阶级、阶层服务；语言是音义结合的符号系统，声音是语言的物质外壳，是发音器官发出的，具有物质性；语言是约定俗成的，文化是人为的；人的语言能力可能是天赋的或遗传的，而文化只能通过后天学习才能继承；文化可以有模式的转换，如政治制度的转变，宗教的转变，社会组织的转变等，而语言的发展是渐变的。从这个角度看，语言是文化中的一个特殊组合体，也可以说语言和文化的关系是局部和整体的关系。

换一个角度看，语言是许多精神文化的载体，是文化的重要表达形式，没有语言许多文化无法表现，如文学、法律、制度、政治、教育等；另一方面，没有文化也不可能产生语言，语言的意义是文化的。可见，语言和文化具有不可分割的关系。从语言和文化的发展来看，两者是相互促进的，文化的发展会推动语言的发展，新词新语的出现、外来词的借入都跟文化有直接的关系；语言的发展也会促动文化的发展，如白话文跟白话文学的关系就极为密切。

不过，应该看到语言和文化的差异，不能把语言和文化看成是因果关系，也不能把语言和文化等同起来。语言作为文化的表现形式，毕竟不同于文化本身，语言有自身的系统和独立的发展轨迹。研究语言和文化的关系主要包括如下内容：① 从语言上观察文化特征，如文化类型、文化影响、文化遗迹、文化演变等；② 从文化上去解释语言现象，如语言的分化和统一、

语言接触和影响、语言转用、语言兼用、混合语、词汇和语言的变化等；③研究语言与文化的内在联系及其发展上的相互联系。需要指出的是，语言可以因不同目的从不同角度去研究，从文化角度去考察语言，只是语言研究和语言解释的一个角度，它绝不是语言研究的全部内容。

语言和文化的关系表现在多个方面，下文仅举某些现象和问题来说明语言和文化有密切关系。

(2) 禁忌语和吉祥语

语言中的词跟其所指的实物从理论上说是没有必然关系的，但是人们常常把语言中的词跟事物联系起来，迷信语言的魔力，这在原始人那里表现得十分明显，原始人认为一些事物不能出现，不便出现，那么表达这些事物的词语也应避免说出来。这些不能说或不便说的词语便是禁忌语，禁忌语也叫塔布(taboo，linguistic taboo)。在多数语言中，涉及死亡、性和生殖、排泄等内容的词语往往在禁忌之列，如英语的“shit(大便)”一般不会被说。被禁忌的事物有时用相应的委婉语表达出来。

在汉语中，“死”这个词是人们最为忌讳说出的词，因而从古到今便出现了许许多多有关“死”的同义委婉语，即便在现代汉语中也有许多，如“病故、逝世、去世、永别、诀别、身亡、作古、夭折、心脏停止了跳动、百年之后”等是书面语中常用的表达方式，口语中可以说“走了、老了、咽气了、永远离开了、过去了、不在了”，如果是坏人死了，可以说“断气了、见鬼去了、见阎王了、到阎王那去报到了”，如果是因公而死或者是党政干部死了，可以说“牺牲、倒下、殉职、献出了宝贵的生命、光荣了”，信仰基督教或佛教的人，可以说“见到了上帝、上西天、驾返瑶池、驾鹤西游、到西方极乐世界”。中国人爱把死说成“见某某去”或“到某某那儿去报到”，如“见马克思/上帝/阎王/见爹妈(已去世)/丈夫(已去世)”等。各地还有许多跟“死”谐音的字眼也在禁忌之列，如不能给老年人送钟表，忌讳送钟跟“送终”谐音。海边的人尤其是渔民忌讳说“住、沉、翻”，这都会引起死亡，所以因“箸”跟“住”同音，改为“筷”，“筷”跟“快”同音；遇到一些东西需要翻过来，不说“翻”，而说“划过来”或“顺过来”。一些地方过年过节害怕小孩子不注意会说出“死”这个字眼，在过年过节前，让小孩子啃猪尾巴，意思小孩子像小猪，不懂事，希望阎王不要在意。上海人去看望病人，不送苹果，因为“苹果”跟“病故”谐音。

不少民族都有对数字的禁忌，这跟对“死亡”禁忌也直接相关。如民间认为 66 岁、73 岁、84 岁都是人生中的“槛”，民间流传着这样的说法“66 不

死也得掉块肉”,“73、84,阎王不叫自己去”,因而,老年人往往忌讳说这3个岁数,常常多报一岁,或少说一岁,以回避所谓的“槛”。中国多数地方忌讳“四”这个数字,因为“四”跟“死”谐音。对“四”的禁忌在日本可谓登峰造极,在日本文部省发布的常用汉字修订本中,汉字“四”被删除,至于删除的原因是因为,在中国、日本、海外华人、海外日本人,这些认为“四”是不吉利的人那里,每个月4日、14日、24日的死亡率尤其是心脏病死亡率要远远高于对“四”没有禁忌的其他民族的人①。电话号码、车牌号码尾号是“4”的,人们也不喜欢。

由于棺材、死人穿的衣裳跟“死”有关,人们就用一个好字眼来代替,于是有了“寿木”、“寿材”、“长生”、“寿衣”等说法。医院里的停尸房叫“太平间”,这也是用相反的意思来命名,以求逢凶化吉。

除了“死”的禁忌语外,社会上还有许多禁忌语。如上海新娘子上门忌讳吃瓜,因为“瓜”跟“寡(寡妇)”谐音。一些地方,对“药”也忌讳,因为“药”总是跟病相关,安徽省定远县把中药叫“香茶”,潮州则把药叫“甘茶”或“利市”,江西一些地方把吃药叫“喝好茶”,湖北一些地方叫“喝细茶”。由于“梨”跟“离(分离、离婚)”谐音,许多地方,尤其是情人、夫妻在分别时不吃“梨子”,梨也不能分着吃,因为“分梨”跟“分离”谐音。

以上的一些禁忌语多数跟迷信有关。另有一些禁忌语是跟不便言说的事物有关,如性器官或性行为。“老二”、“鸟”、“那个”等往往是口语代指男性性器官,写文章时,往往用隐讳号“××”来表示。至于性行为,也有许多委婉说法,如“同房”、“同床”、“房事”、“发生了(肉体、不正当、男女)关系”,文学作品中还常用“云雨”、“风月”、“颠鸾倒凤”、“知晓人事”等说法,“春”字也跟情爱、性有关,如“春心”、“春情”、“怀春”、“春药”、“春宫图”等。不便说的词语往往被认为是不雅的,如改厕所、茅厕叫卫生间、盥洗室,上厕所叫洗手,或者叫方便一下;恋爱、结婚本无不雅,但说出来感到害羞,所以叫个人问题。

禁忌语的出现跟原始文化、原始思维有关。现代人的禁忌语使用也跟文化因素直接相关,如精神安慰,避免粗俗等。避免出言不逊,有伤风雅,斯文扫地是现代人使用禁忌语的主要目的。

不能说,不便说的需要禁忌,而另一方面为了祈求吉祥如意,又要说一些吉利的话。吉利的、好听的话就叫吉祥语。使用吉祥语也跟文化有关,一

① 《中国商报》2003年4月24日。

般人认为吉祥语具有逢凶化吉、富贵降生的神秘莫测的力量，这也是语言崇拜的一种表现。大部分吉祥语是通过实物名称跟语言里表示吉祥喜庆的字眼谐音构成的，所以，同音或谐音是构成吉祥语的主要手段。如“8”在某些方言中跟“发”谐音，“发”寓意发财、发达、发迹，所以“8”就成了一个吉祥的数字，选日子要选初八或8号、十八或18号、二十八或28号，8月8号或八月初八是最好的日子，挑房号最好是808，电话号码、车牌号最好尾数是8。桔子的“桔”在许多地方由于跟“吉祥”的“吉”谐音，许多人家每逢年节总买一盆金桔放到家里，以求吉祥如意。春节时常常在门上贴一个倒着的“福”字，寓意“福到”。安徽的一些乡村，年夜饭中必有一条大鱼（最好是鲤鱼），而且不准吃，年初一以后才可以吃，这样会“年年有鱼”，寓意“年年有余”。安徽定远人春节前后爱送一种方片糕（糯米做的，因为是白的也叫云片糕），寓意高高在上，高兴的意思，出远门的人也送方片糕，寓意步步高，考取大学，也送，寓意“高中（zhòng）”。一些地方，结婚时，要在新房里放上枣子、花生、桂圆、百合，寓意早生贵子、花着生（有男有女）、高贵圆满、百年好合。一些地方还有结婚吃石榴或送石榴的风俗，寓意多子多福。

说吉祥语在汉语中表现多样，说吉利的话，取吉利的名字，选吉利的日子，寄托着一种对美好生活的期盼。这多数是一种民间风俗习惯，图个吉利美满，也不必都看成迷信。

（3）称谓与文化

所谓“称谓”，是“人们由于亲属和别方面的相互关系，以及身份、职业等而得来的名称，如父亲、师傅、厂长等”①。可见，称谓反映出人与人之间的不同关系。称谓可以分为两大类，一是亲属称谓，一是非亲属称谓。

亲属称谓是亲属制度的语言形式，它以简单的名称形式反映出复杂的亲属制度，由于亲属制度反映了婚姻家庭中人与人的关系，是人类社会的重要文化因素，因而亲属称谓跟文化的关系极为密切。不同的民族有不同的亲属制度，因而在亲属称谓上会有不同程度的差异，如称谓的繁简、义项的多少、语义的不同关系、范畴的不同等。如夏威夷的玻利尼亚人的亲属称谓保留了血缘家庭亲属制②的特点，其特点是：① 母方亲属和父方亲属不分；

① 《现代汉语词典》2002年增补本。

② 血缘家庭，血缘婚，是原始社会的一种婚姻形式，属于群婚的初级阶段，同辈男女互为夫妻，不同辈分之间禁止性关系。

② 同一辈分不分亲疏，称呼相同；③ 祖父母以上的辈分都称祖父母，孙以下的辈分都称孙。这种亲属制度只有 5 个范畴：祖父母、父母、兄弟姐妹、子女、孙子(女)。在我国的景颇族的亲属称谓也保留了血缘婚(尽管这种婚姻制度在景颇族中现在已不存在)的特点，如亲兄弟、堂兄弟、姨表兄弟同称；连襟以兄弟相称，父亲的连襟兄弟称父亲，祖父的连襟兄弟称祖父；妯娌间以姐妹相称，母亲妯娌称母亲，祖母的妯娌称祖母；祖父母和外祖父母同称，孙子和外孙同称；岳父、舅父同称，岳母、舅母同称，公公和姑父同称，婆婆和姑母同称；有时不同辈分同一个称呼，如儿媳和弟媳妇同称。血缘婚使称谓词往往具有语义的广泛性，亲疏称呼相同，不同辈分有时没有严格界限。有的民族同一个称谓词可以指称多个辈分的人，如印第安塞米诺语可以用一个词指 7 种人：父亲、父亲的兄弟、父亲的姊妹的儿子、父亲的母亲的姊妹的儿子、父亲的姊妹的女儿的儿子、父亲的父亲的兄弟的儿子的儿子、父亲的父亲的姊妹的儿子的儿子。可见，亲属称谓实际上是亲属制度的反映。

在中国，亲属称谓十分复杂，也十分严格。在我国长期的封建社会中，家庭是社会组织的基本单元，有着严格的宗法制度。这反映在亲属称谓上，是长幼有别，嫡庶有别，亲疏有别，内外有别，系统分明，层次清楚。如按家庭派系正支(妻及其子女)和旁支(妾及其子女)不同；如按血统远近分内亲和外亲，内亲中如妻子的弟弟叫内弟，妻子兄弟的儿子叫内侄，外亲中母亲的父亲、母亲叫外公、外婆，姐妹的儿女叫外甥、外甥女，女儿的儿女叫外孙、外孙女；如按性别分，女系有婶母、伯母、姑母、舅母、姨母，男系有叔父、伯父、姑父、舅父、姨夫，其中叔父和伯父的子女是堂兄妹、堂姐弟，而姑父、姨夫、舅父的子女属于表兄妹、表姐弟，表系中又分姨表和姑表。直系亲属中，如晚辈之子、孙、曾孙(重孙)、玄孙、来孙直至云孙，共 8 代，长辈如父、祖父、曾祖父等，代代清楚明白，界限严格。汉语中这些亲属称谓如果搞错，就叫乱了辈分或乱了套了。

而在西方的称谓中如英语，称谓分别有时不甚严格，如现代英语中，“侄”和“甥”不分，都是 nephew 或 niece；“姨母”、“婶母”、“伯母”、“姑母”、“舅母”用一个词 aunt，“叔父”、“姨夫”、“舅父”、“伯父”、“姑父”用一个词 uncle。这说明，西方人宗法观念比淡薄，宗法等级关系不太严格。

当然，随着中国独生子女生育制度的实行，必定带来家庭结构和亲缘关系的变化，未来中国的亲属称谓必定简化许多。

非亲属称谓也叫社交称谓，指社会交往中所应用的称谓词语。在汉语

中,社交称谓有特殊的表达方式。首先社交称谓中的通称词会随着社会文化的变迁而变化。如20世纪50年代、60年代、70年代的社会通用称谓是"同志",带有一定的政治色彩;80年代流行的是"师傅",师傅本指传授技艺的人或对有技艺的人的尊称,作为社交通称词,也暗示了对称谓对象的尊敬;90年代以后,流行的是"先生"和"小姐"(城市或发达地区),先生和小姐在汉语中含有有地位、有文化、有知识的社会文化意义,作为社交通称词也有尊敬的意义。师傅、先生、小姐相对于"同志"来说,最大的变化是通称词失却了政治色彩,而把对人的尊重放在首位。目前,一些地方还流行把"老板"作为通称词,这是商品经济的反映。在背称的语境下,研究生称导师,课题组成员称负责人,职工称单位领导,乃至一般事业单位、政府机构的工作人员称党政领导,也叫"老板"。有意思的是"小姐"一词命运的变迁,跟社会文化直接相关。在50年代以前,"小姐"指有钱人家的女儿,属于尊称;50年代到70年代,"小姐"成了贬义词,如"资产阶级小姐"、"小资产阶级小姐"、"娇小姐"等;90年代"小姐"成了对一般年轻女子的尊称;而目前在某些地方,小姐又成了坏字眼,某些地方把从事不正当职业的女性称为"小姐","小姐"成了坏女孩的专门称呼。

汉语的社交称谓中还流行表职称谓,一般结构是"姓+职务",如张局长、李主任、王处长等。表职称谓中遇到职务带"副"字的一般去掉,以示尊重。表职称谓是官本位意识的体现。最近,影视剧中又流行把"长"字去掉,如"黎局长"简称"黎局",还如张处、王队,这种称呼显示领导跟群众或其他人之间的一种亲密关系。

另外,汉语的称谓中尤其在文化层次相对较低的人那里,还流行泛亲属称谓,即用亲属称谓来泛称非亲属,如"叔叔、阿姨"快成了孩子嘴上的通称词。农村还有一种比照孩子的泛亲属称谓,"孩子他叔"、"小宝他大爷"、"他王奶奶"(亲属称谓中也存在,如"孩子他爹")。泛亲属称谓由于把对方看成自己的家人,表示了对对方的尊重。称名(姓)法在称谓中也常使用。称名要注意,一般只用于年长的称呼晚辈,至少应是平辈,否则是不礼貌的,一般也只称双名(恋人之间称单名是亲昵的表现)。称姓法往往要在姓前加"老"以示尊重,或加"小"以示亲切。但加"老"或加"小",很有讲究,女性的前面一般不加"老"。

(4) 语言与社会生活的变异

语言中跟社会生活最密切的要素是词汇,词汇反映社会生活最为敏感,

最为积极，新词的产生、旧词的消亡、词义的演变都跟社会生活有直接的关系。

语言的词汇记录下了历史上曾经出现的事物。一些事物和现象现在不存在了，但被保留在语言中的词汇，通过词汇可以帮助寻找旧事物的影响，可以帮助推断某一时期的社会生活。如郭沫若就利用甲骨文，考证出殷商时代，先民已经从渔猎进入田猎，因为卜辞中田猎有 186 条，渔猎只有 11 条，田猎的对象又以鹿为主，并且野马、野羊、野象也是田猎的重要对象，从而说明那时的黄河流域中部是未经开辟的地方。到了诗经时代，则畜牧业受到重视，证明是《诗经》中对马、猪、羊的分类和名称特别复杂，《诗经·鲁颂·駉》中马的名字就有 16 种，那时按马的颜色、年龄、雌雄，各有不同的名称。

先秦两汉时代，对丝织物的分类也特别细致，农具的名称也很多，这说明古代纺织业、农业也特别发达。

再如，"死"这种行为，不同等级的人有不同的叫法，如天子死叫崩，诸侯死叫薨，大夫死叫卒，士死叫不禄。不同身份也有不同叫法，如父死叫考，母死叫妣，妻死叫殡。这说明古代封建社会等级森严，也说明那时对丧葬礼仪的重视。

从文字上，也可以看出古代社会的一些特征，如《说文解字》是东汉时的作品，既反映了它以前的社会历史，也反映了它那个时代的社会历史。一般说，古代妇女地位低下，如《说文》："妇，服也，主服事人者也。"《说文》中收有许多"女"字旁的所谓坏字眼，如"妄、婪、奸、婢、姘"等。但《说文》收有"女"字旁的字 258 个，许多姓氏是"女"字旁，如"姜（神农氏）、姬（黄帝）、姚（虞舜）"，"妙、姿、姝、娇、姣"等也是好字眼。实际上，在远古时代，妇女在母系社会地位是崇高的，只有到了父系社会及其以后，男子的地位才提高。

新词新语的产生跟社会生活关系密切。新事物，新现象，新概念的出现，必然要反映到语言中来，而且首先是反映到词汇中来，其中新词新语的出现就是社会生活变化在语言中的最明显表现。据统计，20 世纪 80 年代以来，汉语中每年约出现 1 000 个新词语，收入《新华新词语词典》中，仅仅反映 90 年代以来出现的新词语就有 4 000 多条。该词典中收录带"网"字的新词语就有几十个，如：网上录取、网友、网卡、网民、网页、网虫、网关、网址、网吧、网迷、网格、网校、网恋、网站、网教、网龄、网管、网德、互联网、网络文学、网络电话、网络银行……这些新词语的出现，正是对网络时代的反映。

新事物出现除了创造新词语来表现外，还可以旧瓶装新酒，从而引起词

义的扩大或转移。可见，词义的变化跟社会生活的变化也有密切关系，《现代汉语词典》2002 年增补本就在附录中收录了许多词义在近几年中扩大的词，如：包装、保安、菜单、超生、出局、底线、非礼、激活、家教、家居、界面、经典、开局等。

词汇中的外来词也是社会生活在语言中的直接表现。近年来，新词语中外来词占相当比例，科技、教育、政治、经济、文化、日常生活各个领域，都有相当数量的外来词。近年来在数量、领域、传播速度、构成形式等方面都较过去有很大发展，这正是社会生活的急剧变化的外在反映。外来词是社会生活的一面镜子①。

3. 语言与社会阶层

社会阶层是一个含义宽泛的概念，可以指一个民族或一个地区内因阶级、团体、行业、收入等不同所形成的具有一定共同社会特征的人群。

语言虽然没有阶级性，但阶级对语言的影响还是存在的。关于阶级的划分，西方社会学中流行德国社会学家马克斯韦博对社会分层的理论，这种理论按人们的财产、地位和声誉等综合指标把社会成员分成 6 个阶层：上上阶级（老财阀）、下上阶级（新财阀）、上中阶级（高等白领）、下中阶级（普通白领）、上下阶级（蓝领阶级）、下下阶级（体力劳动者）。这种关于阶级的划分并不是马克思主义真正意义上的“阶级”，只不过反映出人们在职业性质、文化程度和生活方式上的一定共同性。但这种划分，作为区分社会方言的阶层标准还是可取的。不同的社会阶层在语言上有不同的表现，形成不同的阶层方言。比如，人们调查发现，英国伯明翰、利物浦的“工人阶级方言”（主要是下层阶级）的特征是辅音的舌根化色彩，诺里奇市的“工人阶级方言”的特征是高喉音，而这些特征是这 3 个城市的“中产阶级方言”所没有的。再如在藏语中，贵族上层人士创造了一些专为本阶级所使用的特殊用词，即藏语中“敬语”，有的敬语跟相应的普通词语形式完全不同，如“头”，普通词语是 $ko^{12}$，敬语形式是 $wu^{54}$。我国封建社会中阶级等级森严，阶级分化明显，具有阶级等级观念的词语就比较多，如“妻子”，天子之妻称“后”，诸侯之妻称“夫人”，大夫之妻称“孺人”，士之妻称“妇人”，庶人之妻才称“妻”。

我国目前的社会阶层还没有明显的区分，不过按照人们的职业性质、文化程度、居住环境等的不同还是有一定的社群特征的。这些社群特征在语

① 参阅陈昌来.《外来词，社会生活的一面镜子》，《语文论丛》（七），上海教育出版社，2001 年。

言上形成了一定的语言变异。如在大中学生、年轻人中流行所谓的时尚词语,或叫流行语,如"爽"、"酷"、"Q"、"飘族"、"丁克族"、"SOHO族"、"小资"、"中产"、"波波"一类的词语,应该是城市有文化的年轻人的用语,所谓"学生腔"中就有许多时尚词语。"网民"是一个新的特殊社会群体,网络流行语就是一种特殊的社会方言,不上网的人对"大虾"、"菜鸟"、"恐龙"、"眉眉"、"斑竹"、"烘培鸡"、"瘟都死"、"酱紫"、"打铁"、"灌水"一类词语的含义恐怕难以理解,对"184"、"687"、"520"、"58006"等数字以及某些符号的含义恐怕也是一头雾水。再如,"先生"、"小姐"作为称谓中的通称词,主要用于大中城市,有一定文化的人更常用,农村较少用或不用;农村人中或年纪大的文化水平不高的城市人(尤其北方城市)中,善于用"大爷"、"大妈"、"大伯"、"大娘"、"大叔"、"大婶"等泛亲属称谓作为通称词。再如打招呼用"你好",在农村也很少听见,多数人还是以现场的实际情景或行为来打招呼,如吃饭前后说"吃了吗?",上班时问"上班吗?",下班时问"下班了?",在菜场说"买菜呀!",等等。

职业或行业对语言有相当的影响,平常所谓"干部腔"、"老师腔"等,实际上就是对不同职业语言习惯的概括。

语言学文献中的"行业语"是不同行业的特点在语言上的外在体现。行业语也叫行话。不同的行业有一些专门的词语,如饮食行业有"红案"(主要指做菜肴的)、"白案"(主要指做主食的)等,煤矿工人有"溜子"(井下运送煤的铁皮槽)、"冒顶"(井壁或顶板的坍塌)等,商业上有"盘点"、"脱销"、"热销"、"热卖"、"促销"、"清仓"、"抢手"、"甩卖"、"脱手"、"营销"、"推介"、"炒卖"、"大卖场"、"震撼价"、"跳楼价"、"特卖"、"营收"、"销势"、"脱销"、"俏销"、"卖点"等,股市上有"上市"、"退市"、"熊市"、"牛市"、"猴市"、"虎市"、"尾市"、"前市"、"后市"、"低价圈"、"大盘"、"割肉"、"主板市场"、"止跌"、"涨停板"、"跌停板"、"停牌"、"债转股"、"原始股"、"问题股"、"期股"、"配股"、"参股"、"垃圾股"、"套牢"、"解套"、"上证综合指数"、"飘红"、"飘绿"等。广义的行业语还包括专科词语,即科技术语。

行话中有一种是故意说得让外行人听不懂的话。这种行话往往是某种隐秘社会群体如帮会组织、特殊行业、流氓集团、盗窃集团、黑社会的联络工具或特殊词汇,所以又叫隐语、黑话、切口、春点、杂话、市语、方语、密语、暗语、秘密语等。早期帮会,如洪门帮就有许多隐语,洋伞叫"洪头"、"独脚"、"乌云",本国伞则叫"开花子"、"遮漏子"、"雨淋子",枪叫"手统子"、"牲口"、

"叫驴"、"喷筒"、"喷筒子",手提机关枪叫"快上快"、"麻蜂窝",盒子枪叫"牛腿子"、"蹄子",子弹叫"种子"、"米子"、"钉子",洋火(火柴)叫"绷星子"、"散花子"、"红光子",香烟叫"云条子"、"三寸子"。如有些地方的偷盗者把警察叫"雷子",派出所叫"红灯",入室盗窃叫"查户口"或"钳工"等。香港黑社会把警察叫做"花腰",打架叫"开片",吸鸦片叫"摆横"。一些行业为了谋生,保护自身利益,不得不创造一些暗语。如旧时当铺数目字不便让外人知道,其秘密语中就把数目字"1、2、3、4、5、6、7、8、9、10"分别用"由、中、人、工、大、王、夫、井、羊、非"来代替,以上下左右笔画露头多少为标志。山西晋南把行业秘密语叫"言话",如理发行业中,把理发叫"扫苗儿",烫发叫"烧苗",常客叫"老点",收钱叫"卡把",钞票叫"出儿"。乞丐在中国历代社会中存在,因而很容易形成地下社会组织,即所谓的"丐帮","丐帮"中不同乞讨方式往往有不同的名称,如:

告地状:用韧性很强的纸,写上凄苦的言辞来诉说自己的不幸。

脱轴头:又称诉冤党,假装访亲不遇,向人哭诉苦情,以女丐为多。

观音党:几个妇女纠合在一起行乞。带上孩子的叫"送子观音",一边诉苦,一边让孩子大哭。

地黄党:男丐装扮成双腿残疾,在地上滚行。

拐仙:男丐装扮成单腿残疾,拄杖乞讨。

画仙:男丐装扮成哑巴,指手画脚行乞。

争把子:硬讨钞票的乞丐,苏沪一带称作"强叫化子"。

卖野人头:以恶劣的物品卖高价。

文相夫:用名片、工作证、介绍信行乞。

磨街石:伏在地上行乞。

黑话或隐语是极为特殊的社会方言,随着某个特殊社会群体或行业的消失或公开,某种黑话或隐语就可能消失,如旧时的洪门帮、青帮消失了,其帮会的黑话也就消失了。旧时东北土匪有许多黑话作为联络工具,小说《林海雪原》中就记录了一些,现在也消失了。新的特殊社会群体出现,也会产生一些新的黑话,如近年来地下赌博、盗窃、贩毒、吸毒、走私、偷渡、色情、传销等非法组织或行业,都可能产生一些新的黑话。

#### 4. 语言和性别、年龄

(1) 语言和性别差异

语言的社会变异还体现在性别和年龄上。就性别来看,使用同一语言

或方言的男女之间往往有一定的差异。这种差异在语音、词汇、语法结构以及整体风格上都有表现。如语音上，北京有些青年女性或自以为年轻的中年女性，常把 j、q、x 发成 z、c、s，形成所谓“女国音”。而男青年中很少这样，如果这样则被讽刺为“娘娘腔”。北京的男青少年说话时轻声音节特多，而且常常把舌尖辅音发成卷舌元音，如“反正”好像“反二”，“保证”好像“保二”，这在女青年少见。在美国英语中，妇女发高元音比男人更高，发低元音比男人更低，发前元音比男人更前，发后元音比男人更后。

语言的性别差异在词汇上的表现更为明显。如我国封建社会中，男人自称可以用“我”、“吾”，妇女则一般不用，而用“妾”、“奴”、“奴家”等。在现代北京口语中，“盖”、“盖帽儿”、“派”、“份儿”、“绝”、“没治”、“混蛋”、“他妈的”等词语多为男青年所用，女青年少用；“讨厌”、“缺德”、“不得好死”、“该死的”多为女性所用。北京话中，女性在疑问句中使用“吗”、“呢”、“吧”、“啊”的频率远远高于男性①。像英语中，某些颜色词是妇女专用的，如：beige（米色）、ecru（淡褐色）、aguamarine（蓝绿色）、lavender（淡紫色）、mauve（紫红色），除非从事室内装饰的专业人员，一般男性如果使用这些词是会被嘲笑的。另外，一些形容词也属于女性专用词，如：adorable（极可爱的）、charming（迷人的）、sweet（甜蜜的）、lovely（可爱的）、divine（好透了）等。

语法方面，女性在句类选择上更喜欢用疑问句、祈使句，说英语的妇女比男子更多地使用附加问句。在语言风格上，有的语言学家甚至提出女性语体（female style）的概念，如美国语言学家 R. Lakff 指出，女性语体要比男性语体委婉、犹豫、含混。具体表现在下面几个方面：① 女性更多地使用附加疑问句；② 经常用闪避词（hedge），如“sort of”，“I guess”，“I think”；③ 经常使用强化词（intensifier），如：so, very, really, absolutely 等；④ 更多地使用礼貌用语，如：please, thank you, you are so kind 等；⑤ 使用一些专用颜色词和形容词（见上文）。

在我国湖南省江永县、道县一带还发掘了一种瑶族妇女的专用文字，用来书写湖南汉族人的土语，即所谓“女书”、“女字”。这套文字约有 1 000 多个单字，代表当地土话的 300 多个音节。这种女书虽是方块汉字的变异，字形呈长菱形，由右上向左下倾斜，却是一种单音节音符字表音文字。过去，当地瑶族农家妇女多用七言诗形式将自己的苦难身世，结交姊妹的情谊，祈

① 曹志耘.《语气词运用的性别差异》.《语文研究》.1987(3)

祷祝福，乡间趣事，历史事件，唱本故事等内容写在自制的布面本、扇面、帕子、纸片上。

至于造成语言性别差异的原因，社会语言学家提出了如下一些看法，如从妇女所处的社会地位来看，人类进入父系社会以后，女性越来越处于附属的地位，弱者的心态表现在言语上自然是比较犹豫、委婉，讲究礼貌。另一方面，社会对女性的歧视，使得女性更加敏感，并逐渐培养出一种爱慕虚荣的心理，表现在语言上则是呈现出一种娇声娇气、嗲声嗲气的风格，像北京话中的"女国音"，大陆都市的女孩模仿港台腔等都是出于这种原因。

(2) 语言和年龄

同一个民族或社会中，同一种母语或方言中，不同年龄人群的语言差异是比较明显的，最明显的是许多方言有所谓老派和新派的差异，中老年人的话是老派，青少年的话是新派，老派方言的特点是保守、稳定，新派方言的特点是有较多的新生的方言成分，特别是新的词汇。另外，老派方言跟新派方言的明显差异是语音成分或语音系统上的不同。如宁波方言中老派有 6 个或 7 个声调，新派只有 5 个声调，其中阳上并入阳去，阴上并入阴去。又如广东韶关老派方言中有 18 个声母，新派方言中只有 17 个声母，老派方言中，[n]和[l]是两个音位，而新派方言中[n]已经并入[l]；从韵母看，老派方言是 47 个，新派方言是 52 个，但少了 4 个入声韵。韶关话中老年人接近客家方言，年轻人则接近粤方言。上海话也分老派和新派，下表是上海话老派和新派语音上的部分差异：

| 字<br>派别 | 1<br>南 | 2<br>碗 | 3<br>面 | 4<br>雪 | 5<br>脱 |
|---|---|---|---|---|---|
| 老派 | ne | ue | miI | siI? | t'œ? |
| 新派 | nø | uø | mi | ɕiI | t'ə? |

双语或双方言地区，不同年龄段的人情况会有不同，如散居在关内的朝鲜族，祖辈说朝鲜语，父辈操双语——朝鲜语和汉语，孙辈则转用汉语。据 20 世纪 60 年代的调查，黑龙江爱辉县大五家子的满族，50 岁以上的老人都会讲满语，四五十岁的人能操双语——满语和汉语，三四十岁的人基本上说汉语，但能听懂满语，少数还会说，20 岁以下的人就只会讲汉语了。再如云南某些地方的普米族，中老年人除了讲白语和汉语外，还能讲普米语，而年

轻人则基本上转用白语和汉语，普米语已不会说。

从用词上看，用词的差异往往跟年龄有关。据陈松岑等在 20 世纪 80 年代的调查，北京城区两代人使用亲属词去称呼上一辈非亲属时，有很大差别。35 岁以下的青年一代使用亲属词去称呼非亲属长辈的比例大大下降；中老年和青年各有对方基本上不使用的亲属词，如中老年人用来称呼未婚妇女的"姑"和"姨"，年轻人几乎不用，而年轻人常用的"阿姨"，中老年人也几乎不用；"大爷"、"大妈"、"大叔"、"大婶"一类的亲属称谓词在中老年人那儿使用频率很高，而年轻人更多的是使用"叔叔"、"阿姨"。

语言的年龄差异主要是由不同年龄段的人对待事物的态度和心理特征造成的，由于老年人的保守和青年人的喜欢创新，以致两代人的语言产生一定的差异，年轻人具有赶潮流、求创新、崇尚时髦的特点，因而其在语言上的变化要快于老年人，一些时尚词语、时髦的表达方式，总是先在年轻人中出现。

## 参考书目

1. 侯精一主编. 现代汉语方言概论. 上海教育出版社，2002
2. 游汝杰. 汉语方言学导论. 上海教育出版社，2000
3. 彼德·特鲁杰尔. 社会语言学导论. 陕西人民出版社，1990
4. 赫德森. 社会语言学. 中国社会科学出版社，1990
5. 陈　原. 社会语言学. 学林出版社，1983
6. 陈松岑. 社会语言学导论. 北京大学出版社，1985
7. 祝畹瑾. 社会语言学概论. 湖南教育出版社，1992
8. 徐大明等. 当代社会语言学. 中国社会科学出版社，1997
9. 郭　熙. 中国社会语言学. 南京大学出版社，1999
10. 陈建民. 中国语言和中国社会. 广东教育出版社，2001
11. 戴庆厦. 社会语言学教程. 中央民族大学出版社，1996

# 第七章　儿童语言发展

## 第一节　儿童语言发展的阶段与儿童语言获得理论

### 一、儿童语言发展研究概述

#### 1. 儿童语言发展的含义

语言是人类最重要的交际工具和思维工具，也是人区别于其他动物的本质特征之一。在一般人的心目中，一个生理发育正常的人到了一定的年龄就会说话，这似乎是再自然不过的事。研究表明，儿童到了 5 岁左右，他们的语言系统就已基本完善，可以在社会环境中进行基本的语言交际。但对于人是如何学会语言的，语言发展的内在机制和规律是什么，从婴儿呱呱坠地到熟练使用语言来进行交流，其间经历了哪些过程等问题，我们了解得却十分有限。200 多年来，这一系列与人类自身密切相关的问题曾吸引了一代又一代的不同学科领域的研究人员，他们在不同的理论背景下，运用各自的研究方法，试图解开这个对心理学、语言学、人类学、语言教育等学科领域具有科学价值的谜团；研究领域也从最初的对儿童日常语言及行为的记载发展到运用科学的实验方法系统地进行调查与分析，并从理论上进行了深入的探讨。儿童语言发展研究也逐渐摆脱了心理学的附庸地位，发展成为一门横跨多门学科的交叉性学科。

语言发展也称语言获得，指个体对母语的语音、词汇、语义、语法等系统要素以及语言运用技能等的理解和产生的发展过程。对儿童语言发展过程及内在机制的研究，早期是心理学家而不是语言学家感兴趣的话题。

2. 西方的儿童语言发展研究

公元前7世纪古埃及法老关于腓尼基人是最古老的民族,腓尼基语是最古老的语言的“奇妙”的实验,表现了人类对儿童语言的最初兴趣。一般认为,对儿童语言的真正观察和研究始于18世纪后期德国学者蒂德曼的《婴儿行为日记》。此后,达尔文、席格门等学者的著作中也都有关于儿童语言发展的记录。特别是儿童心理学创始人、德国生理学家普莱尔通过对他儿子4岁前的语言发展的记录,于1882年出版了《儿童心理》这部被誉为儿童心理学的伟大的经典著作,引起了人们对儿童语言的广泛兴趣。20世纪初德国心理学大师斯特恩根据他和妻子对他们3个子女的语言发展的观察,相继出版了《儿童语言》和《六岁以前早期儿童心理学》,对儿童语言中的表达性手势、单词句、句法结构等问题作了探讨。

西方早期的儿童语言发展研究只是心理学众多研究领域的一个方面,心理学家以自己子女语言与行为发展的观察与记录作为材料来源,主要目的在于揭示儿童心理发展过程和不同阶段的智力水平,虽然也提出了一些对此前有关儿童语言发展问题的不同观点与新的认识,如普莱尔认为仅仅用“学习”二字是不能完满地解释儿童语言获得现象的,斯特恩则认为儿童有一种言语需要,语言发展是一个自然发展的过程,但是由于缺少大规模的儿童语言发展状况调查,因而早期的儿童语言发展研究缺乏系统性,结论也显得比较粗糙。

20世纪初以美国著名心理学家华生为代表的行为主义心理学诞生,儿童语言研究重新激发了研究者们的兴趣。这一理论经过汤勒曼、赫尔、斯金纳等为代表的新行为主义的发展,在长达半个世纪的时间内,成为心理学界的主导。在此框架内,也产生了多种有关儿童语言习得与发展的理论,如机械性模仿说、选择性模仿说、强化论、中介论、试误说等等。行为主义心理学采用“刺激-反应”模式来解释所有的人和动物的心理现象,包括人类的语言。“语言,不论其复杂性如何,正像我们通常所理解的那样,开始是一种非常简单的行为。实际上语言是一种动作习惯”①。语言习惯是在一系列“刺激-反应”中学到的。这一阶段他们主要关心的是,怎样将儿童语言的发展作为衡量心理发展的指标,如语句的长度、词汇量的增加等。

20世纪50年代以后,心理学和语言学两大学科领域开始携手,共同关

① 华生.行为主义.李维译.浙江教育出版社,1998.219

注语言问题，特别是心理语言学作为一门独立的学科诞生之后，儿童语言研究也从行为主义的禁锢中冲脱出来，研究兴趣从儿童说些什么转向儿童对语言知道些什么。通过对较少的被试者的长期观察，希望了解清楚每一个儿童的言语行为，探索儿童是如何获得一种语言的句法规则的，进而解释儿童语言能力获得的内在机制。

3. 国内的儿童语言发展研究

我国的儿童语言研究起步较晚，五四运动时期，随着西学东渐，一批有识之士开始译介引进西方儿童心理学的研究成果。这一时期，国内的研究主要集中在儿童语言发展的个案追踪记录和对西方有关成果的验证和补充等方面。我国儿童语言研究的真正兴起，是在 20 世纪 70 年代后期。北京师范大学、华东师范大学、北京大学等高校以及相关的研究机构都把儿童语言研究作为研究方向，涌现出一大批高质量的研究成果，出版了一批儿童语言研究专著，发表了大量的科研论文。跨越六五、七五的国家重大规划课题“中国儿童(含青少年)心理发展特点与教育”的立项，表明儿童心理与语言发展研究在我国已受到高度的重视。更为可喜的是，80 年代以后，儿童语言习得与发展的研究吸引了越来越多的语言学家的积极参与，桂诗春、李宇明、周国光、孔令达、李向农等在国外儿童语言研究的成果与方法的介绍、儿童语言的理解与发生、儿童语言语法的发展等方面的研究，大大促进了我国儿童语言研究的发展。

回顾我国的儿童语言研究的发展历程，“在近一个世纪的发展中，由大而化之的概述到深入细致的调查，由事实研究升华到理论探索，由心理学界的孤军奋战到心理、语言两界联袂，由对正常儿童的研究到对特殊儿童的研究，有了较大收获。对于西方和苏联的重要研究成果进行了引进，对于汉族儿童的语言发展有了一个粗线条的描绘，对语言的一些子系统的发展或特定的年龄段的语言发展，正在做较为深入的研究。并提出了一些理论问题。对特殊儿童的语言问题也已着手研究。同国外相比，差距还十分显著。但是，最近的发展势头却令人鼓舞。特别是 1993 年 10 月初，中国心理学会发展心理专业委员会，在北京召开了儿童言语发展协作研讨会，心理学界和语言学界的儿童语言研究者，第一次坐在一起讨论儿童语言的发展问题。这标志着两个学界在共同关心的问题上，将紧密合作，共同推动我国儿童语言学的发展”①。

① 李宇明. 儿童语言的发展. 华中师范大学出版社，1995，22

西方早期的儿童语言研究主要是为揭示儿童的心理发展过程和测试儿童的智力水平等目的服务的，材料也主要来自心理学家自己的子女的日常语言记录，研究观念和方法都带有明显的心理学痕迹。20 世纪 50 年代，随着心理语言学建立，儿童语言研究受到心理学和语言学的双重影响，其研究地位和研究价值得到提高。但是这一阶段的儿童语言研究，主要是为了揭示儿童的心理发展，或是为了验证语言学理论的一些假设。只是到了近一二十年，儿童语言研究才以解决儿童语言发展的问题为目的，表现出独立发展的趋势。研究方法也从早期的个案日记法发展到观察和实验相结合的临床法，再到多种研究方法并用，多角度全方位地揭示儿童语言发展的核心问题。可以说，儿童语言研究在心理学中产生，在心理语言学中发展，到 20 世纪 80 年代以后开始发展为一门心理学和语言学交叉的独立学科。

心理学框架下的儿童语言研究，是从心理学意义上描述儿童语言发展的过程；心灵主义的儿童语言研究，不仅要探讨儿童语言的发展过程，还要研究儿童是如何获得语言能力的以及制约儿童语言获得的因素。经过 200 余年的历史发展，人们对于儿童语言发展的基本面貌已经有了一定的认识。对于制约儿童语言发展的因素以及这些因素的作用方式进行了一些有益的探讨。对于儿童语言发展的内在机制提出了许多假说，形成了一些理论，并探讨了研究儿童语言发展的一些行之有效的研究方法。但是从作为一门独立的学科的高度来看，儿童语言研究的成果还是相当有限的。在未来的发展中，儿童语言发展将作为一门独立的学科向前发展。这是由儿童语言研究的学术积累、学术价值和社会应用价值等所决定的。

## 二、儿童语言发展的阶段和规律

### 1. 儿童语言发展阶段的划分

盼望子女早日成为家庭交际乃至社会交际中的一个成员，是所有父母的期望。新生儿出生时，哭是他们惟一的语言，他们用哭声代表饥饿、疼痛、尿湿等生理反应。而 2～3 个月的婴儿已不由自主地发出“啊”、“咿”、“喔”等声音。语言发育早的，到了 1 岁左右就能有意识地叫“爸爸”、“妈妈”，以后开始用简单的形式表达较多的含义，如用“饭饭”代表“这是饭”或“我要吃饭”等。2～3 岁时儿童发音能力大为改善，词汇也大量增加，已能用单句甚至简单的复合句来表达自己的意愿或周围的情况。言语的发展促进自我意识的萌芽，儿童开始把自己从客体中区分出来，并且会用“我”来表达自己的

需求和愿望。3岁时虽然发音还不完全准确,但他们已基本掌握母语的语音系统。4岁儿童已能清晰地谈话,喜欢与家人及同伴交谈。5～6岁时儿童语言表达能力明显提高,在与成人和同伴的交谈中,以自我为中心的表达逐步减少,能依据别人的言语调整谈话内容,表明这一阶段儿童已基本掌握自己母语的语言系统,能比较自由地驾驭语言这个最重要的交际工具了。

研究表明,儿童语言的发展有阶段性,不同阶段之间是连续而有序的,这种阶段是一个循序渐进的过程,不能颠倒。儿童必须依次走完这些过程,才能真正地掌握语言。从不同民族儿童语言发展情况看,这种阶段性总体上具有普遍性,表现出儿童语言发展的一般规律。但儿童从呱呱坠地到较为熟练地使用语言,大致要经历哪几个阶段以及各阶段持续的时间等,学者们的看法不完全一致,因此不同的学者根据不同的标准和研究目的,将儿童语言的发展划分出不同的阶段。

西方儿童语言学界通常将儿童语言发展的过程划分为前语言阶段和真正言语阶段,其中前语言阶段又分为啼哭阶段、咕咕声阶段、呀呀语阶段以及过渡音阶段。1岁左右儿童开始说出第一批真正的"词",从而进入了儿童语言发展的第二个阶段,这一阶段又分为单词句阶段、双词句阶段、电报句阶段以及完整句阶段等。

许政援等根据数十年对3岁前儿童语言发展情况的追踪调查,认为"3岁前儿童的语言发展是一个连续的、有次序、有规律的过程,不断由量变到质变的过程"。这个过程既有连续性,又有阶段性,据此将3岁前儿童语言发展分为6个阶段: ① 简单发音阶段(出生至3个月)。基本韵母发音较早,但声母很少,主要是h音,有时有m音。② 重复连续音节阶段(4～8个月)。发音增加很多,发重复连续音节,开始有近似词的音,近半岁有的音开始有某种意义。③ 不同连续音节阶段,学话准备期(9个月～1岁)。不同连续音节明显增加;近似词的发音增多;一定的音开始与具体事物相联系,即具有一定的意义,但还没有成为"词"。④ 单词句阶段,正式开始学话(1岁到1岁半)。开始有指向连续音节和近似词的音节增多,后来能说出有一定意义的词,近1岁半有的已能说出少量简单句。⑤ 简单句阶段,掌握最初的言语(1岁半～2岁)。从分析记录到的儿童在这一阶段说出的各类句子的数量比例看:简单句最多,占55%;单词、单词句占37.7%;最简单的复合句只占7.3%。⑥ 复合句开始发展的阶段,掌握最基本的语言(2～3岁)。这一阶段中单词和单词句所占比例明显下降,而复合句上升到30.5%;简单句中复杂谓

语句比例也有上升;另外,句子明显加长,言语表达的内容也有发展。

李宇明(1995)认为儿童语言发展的阶段应根据语言学标准,而合理的语言学标准指的是儿童用什么样的语言形式(语言系统的标准)进行什么样的语言交际(语言运用的标准)。据此,他将儿童语言的发展划分为5个阶段:①声音发展阶段(出生至6个月左右)。这一阶段儿童会因为无条件反射或低级条件反射而发出大量无意义的、非语言学意义的声音。②被动语言交际阶段(六七个月至1岁)。这一阶段儿童仍不会说话,但已对成人的语言有所理解,可以用自己的体态语与成人进行简单的交际。③特殊语言交际阶段(1岁至2岁半)。这一阶段儿童已能发出基本的语音,能主动与成人进行交流。但他们所使用的语言与成人语言差别较大,局限于单词句、双词句以及电报句等儿童的特殊语言。④目标口语发展阶段(2岁半至6岁)。这一阶段儿童已基本掌握成人语言的语音系统和语法规则,具备一定的词汇量和语言运用技能,可以进行一般的日常交际。⑤成熟阶段(6岁至少年时期结束)。这一阶段儿童不断完善自己的语言系统和语言运用技能,他们的语言已与成人语言差别不大。

周国光(1999)从语言单位组合的层级性以及儿童语言结构能力的发展状况的角度,将儿童语言的发展分为3个阶段:词语法阶段、词组语法阶段以及句语法阶段。"如果我们承认,词、词组、句子这三级语法单位之间的差别并不仅仅是单位的大小,而且还包含着结构关系、组合规则以及相应的语法组合能力等方面的因素的话,那么我们就必须承认,在儿童语言习得过程中必然存在着一个词组习得的阶段,这一阶段是儿童在语言习得过程中所不可逾越的。语言构造的层次性决定了由词到词组再到句子是一个逐层组合的过程,这就意味着儿童习得语言的过程中也必然存在着词及其语法规则、词组及其语法规则、句子及其语法规则等体现不同的语法能力的阶段"①。

虽然学者们对儿童语言发展的阶段划分存在着分歧,但都认为儿童语言的发展存在着阶段性,而且这种阶段是连续的。阶段的划分有利更好地认识儿童语言发展的过程,同时也必须认识到,这种阶段的划分并非截然清楚的界限,在一种新的语言现象大量出现以前,其萌芽状态在上一阶段中就已经出现。另外,从不同的角度、出于不同的研究目的,研究者所划分出来的儿童语言发展的阶段不完全相同,即使是相同的阶段,所划定的年龄段也

① 周国光.儿童语言习得理论的若干问题.世界汉语教学.1999,(3):77~83

不会完全一致。

2. 影响儿童语言发展的因素

虽然儿童语言习得与发展作为一门独立的学科时间较短，但经过此前一段相当长的时期的研究，学者们基于不同的理论背景和研究方法，对儿童语言的发展提出了自己的看法。如行为主义者强调模仿和强化在儿童语言发展中的主导作用；心灵主义者则认为人有一种先天遗传的“语言习得装置”；认知主义者认为，语言的认知遵循人类普遍认知发展模式。观点的分歧，促使学者们进一步思考“儿童究竟是如何获得语言的”或者“影响儿童语言发展的因素是什么”这类对认识人类的语言能力具有普遍意义的问题。通过对正常儿童与非正常儿童语言发展过程的调查与分析，学者们普遍认识到，儿童语言的发展是多重因素综合作用的结果，诸如儿童自身的生理因素、儿童所处的语言环境、个体认知发展水平以及个人差异、语言学因素和社会学因素等。

1）不同民族的语言具有普遍性，但相互间也存在着一些个性差异。语言是一套符号系统，是客观事物及其相互之间的关系在人脑中的反映，这种反映对人类来讲具有普遍性，决定了不同民族儿童的语言习得具有普遍性。另外，不同民族的语言在符号系统中对这种现象和关系的反映存在着一定的差异，如词汇因素差异和句法规则差异等，也决定了不同民族儿童的语言习得在共同具有的普遍性的前提下，在某些语言成分的掌握上又体现出一定的特殊性。如对人称代词的掌握，英语和汉语儿童都在 3～4 岁期基本掌握第一人称单数代词，普遍存在着对第一、第二人称的掌握先于第三人称的现象。但由于英语的人称代词有性和格的变化，而汉语没有，因而汉语儿童掌握人称代词比英语儿童早一些，也更容易一些；另外，汉语儿童掌握第一人称代词优于第二、第三人称，第二、第三人称代词之间没有明显的差异；而英语儿童掌握第一、第二人称代词之间没有明显差异，由于英语第三人称代词的词形变化较多，因而掌握起来与第一、第二人称之间的差异比较明显。

2）人脑和语言器官有普遍性，但发育水平有差异。人脑中有关言语的区域以及人的言语器官都具有共性。这个生理因素也影响并决定儿童语言获得的普遍性。但是，儿童言语器官的发育和认知的发展有早晚的不同；两者发展速度的匹配上也有所不同，这就使得儿童语言发展和获得有差异。

3）儿童的认知发展有普遍性，但存在的某些差异能促进或限制语言的获得。儿童获得语言需要以一定的经验和认知发展作基础，并要有一定的

抽象概括水平。儿童的认知是从具体到抽象不断地发展的，因此儿童语言获得也是先具体后抽象、一步步地发展的。但是儿童在认知发展这个方面存在着某些个体差异，这些差异能促进或限制儿童语言的发展和获得。

4）儿童与成人的语言交际、成人的语言教授和儿童相应的模仿学习，以及儿童本身的选择性和主动创造性也是影响儿童语言发展的重要因素之一。许政援等在20世纪80年代曾追踪考察过儿童在11个月到14个月期间第一批词的获得，以此探讨成人言语教授和儿童相应的模仿学习在儿童语言获得中所起的作用。结果表明，不论从儿童所获得的第一批词的数量看，还是从这些词的具体内容看，它们对语言获得都有重要的作用。但是从深入一步的分析中也可以看到：儿童在与成人交际过程中，不是单纯机械地模仿。成人教授的有些词儿童没有模仿，因为儿童模仿时发音器官有困难，又还没有这方面的经验，认知水平达不到；另有一些词，儿童虽然模仿了，但仍没有获得，因为这些词离儿童生活范围比较远。由此可见，成人的言语教授是通过儿童言语器官的成熟、认知的发展以及儿童本身的选择性而起作用的。正是交往过程中成人的言语教授以及儿童的模仿学习及其主动性、创造性都有所不同，造成了儿童语言获得的个别差异。

大量的研究表明，儿童语言发展状况受到多种因素的影响。儿童语言的习得与发展是儿童主体因素与客观环境相互作用的结果。儿童的先天素质状况、生理成熟的程度、认知发展的水平和后天的学习情况在儿童的语言发展中起着十分重要的作用。虽然不同民族甚至同一民族的儿童个体语言发展过程有差异，但同样值得重视的是，不同民族儿童的语言发展具有很强的普遍性，特别是在早期的语言发展中，这种共性表现得尤为明显。

**3. 儿童语言发展的一般规律**

虽然不同民族的儿童在语言习得上表现出一定的特殊性，但正如上文所说，人类语言具有普遍性，人脑和语言器官具有普遍性，人类的认知发展具有普遍性，因而儿童语言的发展表现出一定的共同性：

1）儿童的语言发展具有先天的生物学基础。儿童所以能获得语言，至少部分地是由于人的一些先天的生物学特点。许多别的动物虽然也以某种有规则的方式沟通信息，但是这种沟通活动是有限的，并且难以改变。在所有的动物中，惟独人具有得天独厚的高度发达的交际系统，惟有人能够真正地说话，这也许是因为人具有特殊的适于发声的器官结构，并具有控制声道和处理复杂符号的先天的能力。

儿童的语言发展还与儿童的生理成熟,尤其是与脑的发育及脑功能的完善有关。先天性脑发育不全的婴儿往往显示了低下的语言能力。大量的研究表明,在脑的发展和语言发展之间存在着高度的相关性。此外,视听器官和发声器官的正常与否,也会给儿童的语言发展带来影响。

2) 置身于语言环境是儿童习得语言的必要条件。儿童与生俱来的生物条件只是获得为儿童获得语言提供了基础,但仅凭这一点并不能保证儿童获得语言和发展语言的能力。那些生来就脱离了人类社会的语言环境的儿童不能自发地产生人的语言,而生活在不同语言环境中的儿童却发展了不同的语言。尽管对于学习和强化在语言发展中的作用还有争论,但儿童置身于一种语言环境并学习这种语言,无疑是获得和发展这种语言能力的必要条件。一个从未听说过某种语言的人要掌握这种语言是不可能的。许多聋哑儿童不会讲话的主要原因是聋,他们的发音器官并没有缺陷。

3) 语言发展与认知发展有密切的关系。认知中的语义编码是语言与认知之间联系的桥梁。例如,如果儿童没有发展出时间的概念,要使儿童掌握标志时间的词和时态变化的语法是不可能的。儿童分析和产生语言,是以两种基本的认知能力为前提的;他必须在不会说话时就能够领会父母或其他成人所说的话;他必须注意他们怎样谈论事情。只有具备了这两种能力,他才能通过语言将声音和意义联系起来。所以,儿童必须具有领会意义的某种方法,而且必须具有领会那些与意义相联系的声音是怎样构成的方法。

语言发展与认知发展的密切联系还表现在语言发展与思维发展的关系上。一方面,语言发展为认知发展提供了概念赖以存在的符号;另一方面,思维发展的水平又决定了语言发展的水平。当儿童还处在皮亚杰所说的感知运动阶段和前运算阶段时,儿童只能掌握情境性的语言,而当儿童处于具体运算阶段时,儿童才能掌握连贯性的语言。抽象的词和语法的掌握有赖于认知的发展,而语言的掌握又对认知的发展起推动和加速的作用。

4) 理解先于语言的产生。儿童最初的言语活动是从听懂成人说出的词开始的。1 岁左右的儿童大约能听懂 10～20 个词,但说出的词却很少,即使能说出几个词,也往往不完整。因此,理解能力的发展是先于生成能力的发展的。语言理解先于语言产生不仅在语言发展的开端是如此,在以后的发展中仍是如此。当儿童还听不懂具有某种句法结构的句子时,要生成具有这种结构的句子是不可能的。

5）掌握简单的句子然后才能逐渐掌握和产生复杂的句子；先掌握具体的词，然后才能掌握抽象的词，儿童语法的发展是由不完善到完善的。儿童语言发展的这种趋势同样说明，儿童语言的发展是受认知发展制约的。

## 三、儿童语言获得理论

### 1. 行为主义框架下的后天环境论

这种理论以巴甫洛夫的条件反射和两种信号系统学说以及华生的行为主义学说为理论基础，强调后天环境因素对儿童语言发展的影响，否认或轻视先天因素或遗传因素在儿童语言发展中的作用。行为主义心理学认为，人只有行为，没有精神，人的心理、意识等精神因素最终都应归结为行为。一切行为都可以视为有机体对环境造成的刺激所作出的反应。环境决定一切，遗传不起作用。人类的语言也是一种行为，人学会说话是外界环境的作用，没有遗传的因素。语言这种行为跟人类的其他行为一样，也是通过习惯的养成而学会的。在他们看来，儿童掌握语言，就是在后天环境中获得语言习惯，而语言习惯的形成，则是一系列"刺激(S)-反应(R)"的结果。在长达半个世纪的时间中，行为主义心理学始终占据着统治地位，在不同时期为了弥补前一阶段某种学说的不足，行为主义学者也不断地对他们的学说进行修正，因此先后出现了多种关于儿童语言习得的学说，如模仿说、强化论、中介说、试误说等等。

模仿说认为儿童对成人语言的模仿是他们掌握语言的主要途径，成人的语言是刺激，儿童的模仿是反应。早期的模仿说可称为机械模仿说，这种学说把儿童的语言看作是成人语言的翻版，否认儿童在语言习得过程中的主动性和创造性，这种观点显然是有失偏颇的。模仿在儿童语言习得过程中确实具有重要作用，但也有很多现象表明：第一，语言中的句子数量是无限的，成人不可能将所有的句子都作为样板，让儿童去模仿；第二，儿童能说出成人从未说过的话，如儿童在双词句及电报句阶段所说出的句子，有不少在成人语言中根本不可能出现；第三，儿童受到生理发育程度的限制，成人的有些句子，儿童不能很好地模仿。

为了克服机械模仿说的不足，有学者提出了选择性模仿说。这种学说认为，儿童对成人语言的模仿是有选择性的，主要是对成人语言结构的模仿，而不是对内容的模仿。通过模仿儿童学习和掌握了相关的语言结构框架，在新的情境下，儿童可以在这种框架内填入新的内容，或者将模仿的结

构重新组合，这样就产生了儿童自己的语言。

选择性模仿说注意到了儿童在语言习得过程中的主动性，但是模仿只是儿童语言学习的一种方式，不能对儿童语言的发展作出充分而合理的解释。

20 世纪 40～50 年代在模仿说和"刺激-反应"说的基础上形成了儿童语言发展理论的强化说，代表人物是斯金纳。

斯金纳是行为主义心理学的集大成者，其代表作《言语行为》集中反映了行为主义的语言学习观。斯金纳认为，解释人类言语行为的最合适的模式就是刺激-反应论，特别是工具型的条件反射论，并广泛地运用强化这一概念在解释儿童的语音、词汇、语法的习得过程。如儿童语音发展过程中最初能发出母语中根本不存在的声音，进入音位系统发展阶段后这些音逐步消失。对于这一现象强化论认为，儿童早期发出的声音，是由其父母有区别地加以强化形成的。不属于母语的音逐渐从儿童的发音中消失，而属于母语的声音由于得到父母的强化而保留下来。复杂句子也是通过对接近于复杂句子的不断强化和累积而逐步习得的。后期，斯金纳提出用"强化依随"的概念来解释儿童言语行为的形成过程。所谓"强化依随"包含两层意思：一，他人发出的强化刺激紧跟在儿童言语之后发生，最初被强化的是儿童偶然发生的动作，反应和强化只是一种时间上的先后关系，没有明显的目的性；二，强化是渐进的，当儿童对成人的示范话语模仿得比较相近时，这种模仿就得到了强化，然后再逐步强化为与成人示范句相近的话语。通过这种渐进性的强化过程，儿童最终掌握了较为复杂的句子的结构。

与模仿在儿童语言发展中的作用一样，强化也只是影响儿童语言发展的因素之一，单凭这一因素无法全面地解释儿童语言发展的过程。同样有事实表明，就儿童语言的形式和内容而言，成人或父母更关心儿童语言内容的合适性，而往往对语言的外在形式不太注意。因此对儿童语言发展而言，父母的强化并不总是合适的。另外，一种语言中的句子是无限的，父母不可能通过强化语言中的所有句子让儿童学会语言，因而这种学说与模仿说一样，无法解释儿童语言发展中的创造性言语的发生。

为了弥补传统的刺激-反应论过于简单化的不足，美国心理学家奥斯古德、莫勒等提出了中介论，或称传递论。这种学说认为，语言能力的发展是一系列刺激-反应的结合，言语行为中的一个词、一句话都具有刺激的性质，从而诱发出条件反应。一种反应又产生另一种刺激，再诱发另外一种条件反应，如此循环，形成一系列联想的序列，即在外显的刺激和反应之间，有一

系列因为联想而引起的隐含刺激(传递性刺激)和隐含反应(传递性反应)所构成的中介体系,从而将传统的"刺激-反应"模式改成了"刺激-传递性刺激-传递性反应-反应"的新模式。

中介说试图通过介于外显的刺激与反应之间的传递性刺激和传递性反应来解释客观环境怎样通过语言作用于人、语言怎样表现非当时当地的事物、新的话语怎样创造出来并被理解等传统的刺激-反应论所不能解释的问题;并利用它进而解释儿童是怎样通过一系列的刺激-反应链学会语言的。这是一大进步。但是由于中介说不愿放弃刺激-反应的基本模式,仍有很多地方不能自圆其说。正如当代美国的心理语言学家福多尔所批评的:传递性反应不一定是在刺激的作用下直接产生的。如很多人在没有见过一个小偷之前,就能从间接经验中获知小偷是什么。传递性反应也不一定能够成为隐含的刺激,引起新的反应。比如,如果人们不相信"某人是小偷"说的是事实,就不会对这句话或是对某人的名字产生中介说所说的那种反应。

为了解释儿童语言的发展,桑代克(Thorndike)提出了"试误说",这种理论试图用刺激与反应之间的联结的概念来说明人的学习过程和动物的学习过程。桑代克认为,单纯的重复并不能或很难在刺激与反应之间建立起联系,更重要的是行为的效果,并提出了效果律:

> 在同一情境中产生的几种反应,在其他条件相同的情况下,那些伴随着动物意愿满足的反应便形成了与那个情境的坚固联结,于是当情景再现时,这些反应很可能再现;在其他条件相同的情况下,由于动物意愿不满足相伴随的反应与那个情景形成微弱的联结,于是当情景再现时,这些反应很可能不发生。满足或不满足的程度越大,联结强度就越强或越弱①。

桑代克将对动物学习过程的观察而得出的结论运用在儿童语言学习上,认为学习是在相应的刺激和反应之间建立稳固的联系,儿童语言习得的过程是渐进的"尝试错误"、直至最后成功的过程。与成人的语言相比,儿童的语言充满错误、有待完善,而这种错误是由于儿童语言体系不完善、学习时间不充分、正确的语言习惯尚未养成等因素造成的。

行为主义框架下的儿童语言发展理论由于植根于传统的刺激-反应论,

① 桑代克.动物智慧.现代心理学体系.社会科学出版社,1988.107

忽视了人与动物之间的本质区别，强调通过研究那些可以观察到的言语行为，来建立一套完整的儿童语言发展理论。由机械模仿说到选择性模仿说，再到强化论和中介说，每次发展都充实了一些新的内容，试图以此来克服前期理论的不足。充分肯定了模仿、强化等因素在儿童语言发展中的作用，只不过他们把模仿、强化等视为儿童语言发展的惟一因素或主导因素，因而不断遭致别人的批评。

2. 心灵主义背景下的先天决定论

儿童语言发展中的先天决定论是由乔姆斯基的语言学理论发展出来的。20 世纪 50 年代美国学术思潮开始发生变化，理性主义和心灵主义逐步取代经验主义和行为主义而占据主导地位。行为主义心理学的基本思想是，人的心理、意识之类的精神因素最终都应归结为行为，心理学研究的对象是外在行为而不是内心活动，因此一切研究都要以可以观察到的事实为基础。一切行为都可以看作是有机体对环境造成的刺激所作出的反应，环境决定一切，遗传不起作用；言语也是一种行为，也无非是刺激、反应；言语只有物理形式，没有心理内容，人学会说话完全是或主要是环境的作用，与遗传因素无关。而理性主义认为人的知识除了经验成分以外还有先验成分，心灵主义认为人的行为背后还有心灵活动。由于行为主义重在分析与解释人的外显的语言行为，无法说明人的内在语言能力，乔姆斯基提出了转换生成语法理论，这一理论的提出彻底动摇了行为主义心理学长达半个世纪的统治地位。

与行为主义不同，乔姆斯基将语言看作是无限句子的集合，区分人的语言能力与语言运用，认为不同民族的儿童能在四五年的时间内掌握语言，能说出和理解以前从未说过和听过的话，其根本原因在于人有一种先天的"语言习得装置"。这套装置包括两方面的内容：一是包括若干范畴和规则的语言普遍特征，二是人的先天的评价语言信息的能力。认为儿童获得语言，就是运用先天的评价语言信息的能力，为这套普遍的语言范畴和规则赋上各种具体语言的值。这种学说把儿童语言的发展看作是一个演绎的过程，而非归纳的过程，认为后天环境因素的主要作用在于触发儿童的先天语言习得装置发挥作用。

儿童语言发展中的先天决定论者反对后天环境论者把儿童看作只会对刺激发生反应的被动模仿者，而把儿童获得语言的过程描绘为一个积极主动、充满创造性的过程。认为儿童获得的不是一句一句具体的话语，而是关于语言的一系列规则。这一系列规则能使儿童听懂他从未听过的话，也能

让儿童说出他从未说过的话语。

先天的语言习得装置的活动有一个临界期。过了这个临界期，它就会退化。这可以说明成人学习语言的能力不及儿童，也可以解释脱离人类环境、超过这一临界期的儿童无法习得人类的正常语言的现象。

这一学说的提出大大促进了儿童语言习得研究的发展，注意到了儿童获得语言的先天因素和儿童的主动性、创造性，从根本上改变了行为主义的儿童通过被动模仿习得语言的传统看法，从普通语言学理论的高度来看待儿童语言的发展，具有重要的理论价值。

当这种学说以其行为主义的对立面而引起人们越来越浓厚的兴趣的同时，也有人对此提出了批评。这些批评主要集中在以下几个方面：

1）乔姆斯基所提出的“先天语言习得装置”纯粹是思辨的产物，是一个无法证明的假设。

2）先天决定论对儿童语言发展中后天环境因素的作用没有给予足够的重视，事实上不少研究表明，后天环境因素比如成人对儿童语言的教授、成人的强化对儿童语言的发展具有重要的导向作用。

3）乔姆斯基所说的先天语言习得机制和普遍语法具体包含哪些内容，这一装置是如何运作的等问题，至今仍不十分清楚。

4）先天决定论认为，儿童的语言发展是一个创造性的过程，这种创造性体现为儿童只接触到有限的语言输入便可以习得母语复杂的语言结构，环境因素不能左右和决定儿童的语言发展模式。但这种观点并没有得到有关儿童语言研究长期跟踪调查的统计事实的支持。相反有材料显示，乔姆斯基所说的儿童早期语言习得成果显赫而输入极为有限这种所谓“刺激的贫乏”是没有根据的。

此外，乔姆斯基把儿童语言习得的过程看得过于容易，而有统计材料和研究成果显示，儿童在语言习得过程中投入了大量的时间和精力，这一过程是十分艰巨的，“儿童学习母语或者说第一语言是一个艰苦、复杂而又漫长的过程，而不是像某些人所说的那样是一个轻松愉快而又简短的过程，似乎儿童只需要听到周围的人说过的既不完全又很少的一些话就能凭借天生的语言机制一下子就掌握了这种语言、一下子就什么话都会说了。这种说法是完全不符合事实的”①。

① 胡明扬.语言和语言学习.世界汉语教学.1993(1)

3. 认知主义基础上的先天后天相互作用论

行为主义和心灵主义在儿童语言发展问题上走向两个极端，片面强调后天环境和先天遗传在语言发展过程中的重要性，都不能对此现象做出完满的解释。儿童语言的发展是一个多种因素共同作用的复杂过程，在对行为主义的后天环境论和心灵主义的先天决定论的得失思考之后，学者们以更为冷静的态度重新审视这一现象和过程，提出了儿童语言发展中的先天和后天相互作用论的学说。这种学说以皮亚杰的认知说为理论基础，认为儿童的语言发展是天生的能力与客观的经验相互作用的结果。由于他们用一组认知的程序(或称认知机制)来解释言语行为，因此在心理语言学界被人称为"程序派"。

皮亚杰理论的核心是"发生认识论"，既强调内因外因的相互作用，又强调在这种相互作用的过程中，儿童的心理不断发生着量和质的变化。儿童心理既不是起源于先天的成熟，也不是起源于后天的经验，而是起源于主体的动作。主体通过动作对客体的适应是儿童心理发展的真正原因。在此基础上形成的儿童语言发展理论认为：

1) 人类有一种先天的认知机制，这种机制是一种一般性的加工能力，而不是乔姆斯基所说的由普遍语言的特征所构成。它不仅适用于语言活动，而且也适用于其他一切认知活动。儿童语言发展的普遍性，不是因为他们有一个与生俱来的普遍语法，而是因为人类具有普遍的认知策略。

2) 儿童并没有特殊的语言学习能力，语言学习能力只是人的普遍认知能力的一种。语言不决定认知能力的发展，相反，认知能力的发展决定语言能力的发展。语言能力的发展不能先于认知能力的发展。

3) 儿童的语言发展是儿童主体因素和客观环境因素相互作用的结果，是通过同化和顺应不断地从一个阶段发展到一个新的阶段的过程。

在皮亚杰提出认知发展理论之后，人们不断运用这种理论对儿童语言发展进行研究，并获得了一些新的认识。有学者强调人类存在先天的认知程序机制，并用这种机制来解释儿童语言的发展，因此他们的理论又被称为"认知程序说"。认知程序说吸收了乔姆斯基的先天决定论和皮亚杰的发生认识论的某些观点，又有所发展。其主要观点是：

1) 人类先天有一种认知机制，这种机制适用于一切认知能力，它可以用特殊的方式加工整理语言输入的材料。儿童虽然没有先天的语言知识，但存在着先天的加工语言材料的程序机制。儿童语言习得的基础是先天分

析机制进行加工整理的结果，而不是先天决定论所说的与生俱来就存在的。

2）重视语言运用对语言能力的影响。语言能力是语言运用的前提，但语言运用和语言能力同等重要。

3）环境因素对儿童语言的发展具有重要作用，后天的学习对儿童语言发展必不可少。认为儿童语言习得的过程是从周围环境的可接触的语言输入中不断归纳、提炼语言规则的过程。

基于认知主义的先天后天相互作用论是当前儿童语言学中较有影响的理论。但是这种学说不是专门为解决儿童语言获得提出的，要弄清儿童认知水平的发展与语言结构的发展之间的关系还需要做大量的研究工作。而且这两种发展是否就那样的平行，也还缺乏足够的证据。特别值得注意的是，认知说只强调认知发展对语言发展的影响，而忽视乃至否定语言发展对认知发展的影响，这种"单向"影响的看法，也已经受到越来越多的批评。

纵观儿童语言发展研究中出现的各种理论可以发现，在这一领域，研究者们已不满足于早期对儿童语言的记录、描写与分析，试图从更高层次上对这种现象做出理论说明。同时，种种不同的理论之间虽然存在着差别，但又有内在的联系，从行为主义的后天环境论到心灵主义的先天决定论再到认知主义的先天后天相互作用论，学者们认识到，儿童语言的发展不是某一因素单独作用的结果，而是多种因素协同作用的过程，片面强调某一因素都不能很好地解释儿童语言发展的过程及内在机制。

## 第二节　儿童语音的发展

语音是语言符号的形式，是由人的发音器官发出来的、代表一定意义内容的声音。婴儿离开母体，随着发音器官的发育成熟，能发出大量的母语中有或母语中无的声音，其中有些声音并不代表特定的意义，有些声音虽然能表达一定的意义，但这种意义是宽泛的。这些声音不是严格意义上的语音。儿童语音的发展是儿童语言发展的第一个关键时期，必须经过一个准备阶段，才能发展出完善的语音系统。儿童语音发展的过程可以从声音、语音的听辨和声音、语音的发生两个方面来认识。

**一、儿童声音、语音的听辨**

关于儿童何时具有辨认不同声音的能力，有研究者认为，人的听觉系统的主要器官内耳在母体中就已发育完成，胎儿在母体中就已经具有了听辨不同声音的能力。虽然这一结论还有待进一步证实，但有研究显示婴儿出生不久就表现出对不同声音的反应强度差异，对言语刺激特别敏感。有人发现当成人说话或听话时所产生的短暂的肌肉运动很容易和言语音段同步发生，而出生不到1个月的婴儿也表现出类似的反应。这表明儿童的声音听辨能力是早期发育完善的感知能力之一。康登和桑德把正在听成人话语录音的不到1个月的婴儿的行为拍成电影，然后逐个镜头分析，发现婴儿的身体运动，如头、脚趾、手、胳膊等，在言语开始或停止时也同步地开始或停止。进一步的研究发现，美国儿童的身体行为既和英语同步也和汉语同步，但与其他有节奏的声音不同步。这说明儿童的语音知觉并不局限于他在周围所能听到的声音，关键是婴儿对人类语音与其他声音的反应程度的差异。

研究表明，1个月的婴儿似乎就已具有语音范畴的知觉。7周左右的新生儿已经能区分升调和降调，身体各部位的活动与话语的节奏具有同步性，能分辨出带有不同感情的声音以及清辅音与浊辅音。2～3个月的婴儿开始有意识地注意周围的声音，3个月时能主动寻找声源，4个月的婴儿能区分男声和女声。种种现象表明，婴儿到4～5个月时，已经初步具备了区别声音的音高、音长、音色以及不同的声音所具有的情感色彩的能力。6个月时，能对个别词语做出指令性反应，7个月时可以对少数的祈使句和疑问句做出反应。随后这种带有萌芽状态的对话语的理解的范围逐步扩大，类型也逐渐增多，到1岁前已基本掌握了语调。虽然在这一发展过程中，儿童还不能说出一个真正意义上的词，但他们的听辨能力已先于发声能力而语言化了。

儿童具有了区别语音与其他声音的能力，并不表明此时的儿童已能区分母语中的音位或音素。研究表明音位的辨别比成人语音与其他声音的辨别要困难得多。卡里科曾试图通过实验分析儿童区别清辅音和浊辅音的能力，结果显示儿童直到2岁左右才掌握辅音的清浊的区别，而且当两个无意义的音节具有明显的差别时，较小的儿童已能正确辨别。但是当两个音节的惟一差别是词首的辅音有无浊音化时，他们不能正确辨别。这表明音位的区别不只是辨认的问题，还必须理解在语音中哪些差异是本质的、是和意

义有关的,哪些差异是非本质的。因此儿童区别音位比纯音素的辨认困难得多,发展也相对较晚。

儿童具备辨别语言中不同音位的能力之后,这种音位的差别对儿童来讲其敏感性有先后的顺序。有研究显示,10个月到2岁讲俄语的儿童首先能把俄语中的[a]与词首的其他元音相区分,然后才能把前后元音相区分,把高低元音相区分。这种从粗略到精细的分化过程在辅音的区分中也同样有表现。

音位的感知能力只是儿童语音感知发展的一个方面,儿童还要学习辨别哪些语音的组合是词,哪些不是词。不同的语言中音位的组合有各自的特点,如汉语语音系统中没有复辅音现象,而英语语音中有复辅音,但英语中的复辅音有一定的范围,如"star, skill"中清擦音[s]与清塞音[t]、[k]的组合是英语中允许的,并非任何辅音的连续出现都是英语中的复辅音,如"mvaq"中浊鼻音[m]与浊擦音[v]的组合是英语所不允许的。研究表明,儿童到4岁时才具有辨别词与非词的语音组合的能力。

在儿童语音发展过程中,人们还发现,有些音儿童虽然受到生理发育程度的限制不能发出,但是他们已经能够识别这些音。伯科和布朗发现,一个儿童将他的玩具充气塑料鱼叫做fis(实际发音是fish),但是当成人模仿儿童的发音说"这是你的fis?"时,儿童却一直试图纠正成人对他发音的模仿,直到成人说"那是你的fish?"儿童才说"是的,我的fis。"这表明儿童虽然没能发出[sh],但已经感知到[s]和[sh]的区别。说汉语的儿童也有类似的现象,有研究者发现,儿童把"爸爸"发成"wawa",但当他爸爸让他叫"wawa"时,他却很茫然。这种现象也使研究者认为,儿童记忆中词的语音表象主要是按照成人的发音形式储存的。

就语音的感知而言,儿童由于生理发育等因素的影响,能区别某些他们不能发出的音,表明儿童对语音的感知明显先于他们语音的产生。

### 二、儿童语音发生的准备

儿童从呱呱坠地的第一声啼哭便开始了发声的发展过程。但是1岁以前儿童所发出的音当中,有些没有特定意义的,有些虽然与个人的要求或情绪有关,但这种联系是含糊的、宽泛的,这些音都不能称为严格意义上的语音。但是这些发音与语音的发展有着密切的关系,经过这一准备期,儿童才能产生真正的语音。儿童语音的发生是从他说第一批真正的词开始的,经过一段时间逐步习得母语的语音系统。

儿童语音的准备期从婴儿出生到1岁左右说出第一批真正的词，又可分为非自控音阶段、咕咕声阶段和呀呀语阶段。

1. 非自控音阶段

儿童的第一次发声便是离开母体时的啼哭。这表明儿童的发音器官已初步具有了发出声音的能力。婴儿出生初期，啼哭声是他们惟一的“语言”，这些发音大多是由于外界刺激或内在感觉而产生的本能的生理反应，如饥饿、疼痛、需要帮助等，并非肌体主动控制发音器官发出声音，因此有人称这一阶段的发音为“非自控音”阶段，也有人称之为“生物闹音阶段”或“啼哭阶段”。婴儿出生约1个星期后的啼哭大致可分为3种类型：一是要求得不到满足时的大哭，这种声音响度较高；二是由于疾病或外界刺激引起肌体疼痛或不适而引发的哭声，这种声音短而急促；三是需要帮助的哭声，这种声音有一定的节奏。哭声无论短促还是有一定的节奏，都不是严格的语言学意义上的语音。但也有学者通过观察分析这些哭声和婴儿发出的其他声音，认为这一阶段婴儿所发出的音已能区分出辅音和元音以及不同音的组合，“较早出现的辅音是发音部位靠后的喉音，接着出现双唇音和喉壁音、小舌音；发音方法有不送气塞音、送气塞音、鼻音、浊擦音、搭嘴音、吸气音；浊辅音多清辅音少。元音有不圆唇的低元音和发音部位稍高的前元音。这些辅音和元音构成的音节有c(辅音)、v(元音)、vv和cv 4种类型，其中以c型为最多”①。但无论如何，这些发声需要不同发音器官的参与，声音的变化对协调各发音器官的活动具有一定的作用，而且不同的啼哭声所带有的升降模式，也为后来语调的发展准备了必要的条件。

2. 咕咕声阶段

这一阶段将持续到婴儿5个月左右。这一阶段除啼哭以外，其他的发声行为如笑声开始出现，并且逐渐增多，发声的种类也比前一阶段有所增加。4个月时已能发出比塞音、鼻音等发音难度高的边音以及部分卷舌塞擦音，大部分元音都已出现。另外，这一阶段的儿童能够长时间地发出连续的两三个甚至更多的同类或不同类的元音、辅音，其中有些则类似语言中的音节，而且连续发音也表现出一定的语调模式。

3. 呀呀语阶段

此阶段到儿童9个月时达到高峰，主要是辅音和元音结合体的重复，类

① 李宇明. 儿童语言的发展. 华中师范大学出版社，1995. 59

似于成人语言中的音节,如 gagaga、mamama 等等。这一阶段儿童模仿发音的能力大大增强,一方面不断扩充他们所能发出的音,另一方面受到成人语音的影响,先前所发出的一些在母语中没有的音逐渐消失。在咕咕声阶段的连续发音的基础上,呀呀语的声音高低变化类似于语调。有些发音比较接近语言中的词,这大大激发了父母期盼儿童早日开口说话、进行真正的语言交流的愿望,并不断强化儿童先前所发出的类似词的声音。

这一阶段会持续到开始说话之后很长一段时间,大部分呀呀语的声音包括了儿童早期发声中的类似音。这一阶段后期,婴儿的发音开始具有意义,口形也更为准确,意义有专指化倾向,其中个别音节开始具有比较固定的音调。种种迹象表明,经过了 1 年左右的发音准备期,儿童已开始进入真正的语言发展阶段。

在这一准备期中,儿童语音的发展体现出一些共同的规律:"元音(V)首先出现,然后是'辅音+元音'(CV)的结合,之后又是'元音+辅音+元音'(VCV)或'元音+辅音'(VC),最后是前几种语音组合的复排列(CVCV、VCVC 等)。"[①]儿童语音的发展过程是与其生理器官的发育以及发音器官的协调动作密切相关的,语音出现的先后顺序更多地与发音难度有关,而不是与词义有关。

### 三、儿童音位系统的发展

儿童语音系统的发展首先是能发出真正意义上的语音,然后不断完善自己母语的音位系统。音位是某种语言或方言中能区别意义的语音类型单位。音位分音段(音质)音位和超音段(非音质)音位两种,前者是由语音的音质不同在某种特定的语言或方言中区别意义情况的不同决定的,包括元音音位和辅音音位;后者则是由语音的音高、音强、音长等在不同语言或方言中区别意义情况的不同决定,主要有调位、重位和时位 3 种。

#### 1. 音段音位的发展

目前儿童语言学界对儿童音位系统的发展还缺乏系统的研究,但有人认为儿童音位系统的发展有一定的顺序。雅可布逊认为,儿童语音中的音位对立是按照固定的次序出现的,并提出了儿童语音发展的非连续性假说。认为儿童在呀呀语阶段之前能发出大量的各种不同的声音,这些声音是杂

---

① 周国光、王葆华. 儿童句式发展研究和语言习得理论. 北京语言文化大学出版社,2001. 158

乱无章的、无序的，而且不与特定的意义相联系，因此不具有语言的本质特征。但是当儿童进入单词句阶段后，儿童开始习得有语言价值的、与意义相联系的音位，把自己的发音限制在有限的范围内，因而发音急剧减少，而且这种发展表现出很强的规律性：人类语言中普遍存在的共有的语音先习得，母语中特有的语音后习得。具体表现为：

习得唇音/非唇音的对立先于塞音/擦音的对立；第一个辅音一般是双唇音。

习得鼻音/口腔音的对立先于唇音/齿间音的对立。

前辅音的习得先于后辅音；清辅音的习得先于浊辅音。

塞音、擦音、塞擦音的3组音的习得顺序为：塞音＞擦音＞塞擦音。

元音习得的顺序为：[a]＞[i]＞[u]＞[e]＞其他元音。

鼻元音、边音最后习得。

国内学者吴天敏、许政援、张仁俊、朱曼殊、李宇明等通过对汉族儿童语音发展的研究，总结了汉族儿童语音发展的顺序：

辅音发展的一般顺序是：辅音部位是由前后两端挤向中间，发音方法的习得顺序是鼻音、擦音、塞音先习得。

元音发展的一般顺序是：舌面元音＞舌尖元音，不卷舌元音＞卷舌元音，不圆唇元音＞圆唇元音，低元音＞高元音，前元音＞后元音。

可见，国内学者对汉族儿童元音与辅音发展顺序的观察与雅可布逊的预测基本一致。

从语音准备期向真正的语音发展的进程中，儿童在发音上常表现出一些特殊的发音策略。这些发音策略主要有以下几种：

1）省略词尾辅音。如将 ball 发成[ba]，将 hat 发成[ha]等。在省略词尾辅音时，有时会带来元音的舌位变化。

2）简化复辅音。这有多种表现，如简化词首的复辅音中的第一个辅音，简化词首复辅音的第二个辅音，简化词尾复辅音的第一个辅音，或者简化词尾复辅音的后一个辅音，等等。如将 sky 读成[kai]。

3）略去非重读音节。如在英语中将 banana 读成[nana]。

4）替代。在英语中，许多儿童喜欢用浊辅音替代清辅音，用擦音替代词首的塞音，用塞音替代词尾的擦音。

*2. 语调和声调的发展*

关于儿童语音发展中语调的发展情况已有很多成果，研究表明婴儿期

的发音行为就已经具备了成人的语调模式,有研究显示,5～6个月的婴儿开始运用语调来进行交际,8个月婴儿的发声中就已出现升降调。

由于印欧系语言不采用声调来区别意义,而以往关于儿童语言习得的研究主要以印欧系语言为主,因而有关儿童声调的发展研究相对较少。汉语是声调型语言,声调是汉语音节的构成要素之一,在汉语中具有重要的区别意义的作用。因此对汉族儿童来说,声调的习得是语音发展的一个重要方面。有人认为,汉族儿童声调系统的习得远远早于语音系统的其他方面因素的习得,儿童早期就能模仿母语的声调,对声调的发音控制在习得元音和辅音之前就已经完成。赵元任曾对他的孙女的声调习得情况进行过观察,发现在16个月时,他的孙女把第二声与第三声混为一个声调,都变成第二声,但到2岁时已经完全习得了汉语的4个声调,这种现象在单词的发音中表现特别明显,同时开始出现连读变调。李讷和汤珊笛通过对台湾17个儿童为期7个月的调查发现,汉族儿童在习得元音和辅音之前就已习得了声调,而且习得第二声的难度比习得第一声和第四声的难度要大。这与H·克鲁麦克对汉族儿童声调习得情况的观察基本一致。吴天敏、许政援的研究表明,汉语声调在儿童9个月时的发音中就可以观察到,除阴平调外,其他3个声调都已经出现。

虽然研究者对儿童声调习得顺序的认识不完全一致,但都倾向于认为声调和语调在儿童语言习得的早期已经出现,有的先于语音发生,也有的与语音同时出现。

语音作为语言符号的形式,在儿童语言发展中具有重要意义。儿童的语音发展,要受到生理因素、语言因素和环境因素的多种因素影响,在不同时期这些因素的作用不同,儿童语音发展初期,生理因素的影响较大,其他因素的影响较小,由于生理发育的限制,虽然能区别某些音,但不能发出这些音;随着儿童生理发育的日趋成熟,生理因素的影响会逐渐减弱,而语言因素和环境因素的影响相应地会逐渐增强。

## 第三节　儿童词汇的发展

词汇是语言的结构要素之一,是一种语言中所有的词和固定短语的总

和。人们在运用语言这个工具交流思想的时候，词是最基本的要素。词是形式和内容的统一体，所谓形式，包括外部形式和内部形式，外部形式即每个词都有读音，内部形式即每个词也都有其结构方式。词的内容也包括两个方面，一是词汇意义，一是语法意义。词汇的发展是儿童语言发展中的重要阶段，儿童在经过了语音准备期以后，具有了发出语音的基本能力，在多种因素的影响下，儿童进入了发出真正意义上的词的时期。

### 一、不同年龄段儿童的词汇量

词汇量是衡量儿童语言发展水平和认知发展程度的重要标准。考察儿童在不同年龄段的词汇量主要有两种方法：一是考察儿童在不同年龄段的总词汇量及增长方式；二是研究不同年龄段各类词的比例及其发展情况。但是由于不同的学者考察的对象不同以及研究方法的差异，得出的结果也不完全相同。下表是不同的心理学家所考察的不同年龄段儿童的词汇量情况：

| | 1岁 | 1.5岁 | 2岁 | 2.5岁 | 3岁 | 4岁 | 5岁 | 6岁 |
|---|---|---|---|---|---|---|---|---|
| 斯特恩 | | 100 | 300～400 | | 1 000～1 100 | 1 600 | 2 200 | 2 500～3 000 |
| 史密斯 | 3 | 22 | 272 | 446 | 896 | 1 540 | 2 072 | 2 562 |
| 史慧中等 | | | | | 1 000 | 1 730 | 2 583 | 3 562 |
| 吴天敏等 | | 950(1.5～2) 1 065(2～2.5) 860(2.5～3) | | | | | | |

从上表中可以看出，对3岁以上儿童词汇量的统计差别不太大，但对刚进入真正语言期儿童的词汇量统计则有较大的出入，如1.5岁儿童的词汇量，最少的只有22个词，而最多的将近1 000个。由于环境因素、个性差异等的影响，有人提出了儿童最高词汇量与最低词汇量的概念。如：

| 年龄(岁) | 1.0～1.2 | 1.3～1.5 | 1.6～1.9 | 1.9～1.11 | 2.0～2.2 | 2.3～2.6 | 3.0～4.0 |
|---|---|---|---|---|---|---|---|
| 最高词汇量 | 58 | 232 | 383 | 707 | 1 227 | 1 509 | 2 356 |
| 最低词汇量 | 3 | 4 | 44 | 27 | 45 | 171 | 598 |

词汇量的统计对于说明儿童语言发展的状况有重要意义，从上表可以看出学者们所统计出的不同年龄段儿童词汇量的状况有一定的出入，对于形成这种状况的原因，李宇明(1995)认为：

一是儿童词汇量的发展本身就存在着差异。儿童词汇量的多少与儿童

所处的环境、接受教育的程度以及所学语言的本身都有一定的关系，彪勒所设立的“最高词汇量”与“最低词汇量”就是要反映这种差异。

二是因研究者而造成的差异。首先，就“词汇量”而言，是指儿童能理解的词的总量，还是儿童能说出的词的总量；用什么方法可以准确地测量这两种词汇量。其次，就调查所得材料的处理而言，词与非词如何划界，这对于汉语尤为突出；怎样才算掌握了一个词；儿童临时模仿说出的词、儿童自己造的而目标语言中不存在的词以及误说的词如何处理等等问题，会直接影响到词汇量的统计。

虽然统计数据有差异，但从中也可以发现儿童词汇发展的大致规律：从总的趋势看，随着年龄的增长，儿童词汇量不断增加；3 岁以后儿童词汇增长的速度逐渐减缓。在儿童词汇发展的过程中，3 岁左右是调整增长期，发展速度最快。

## 二、儿童词汇习得过程中各类词的发展状况

儿童词汇量的研究重要在说明不同年龄段儿童所习得的词的总量，衡量儿童词汇发展状况的另一重要尺度是不同年龄段儿童所习得的词汇中各类词的比例情况。由此可以发现不同年龄段儿童词汇的构成情况，有研究显示，从儿童词汇中各类词的发展状况看，实词占有绝对优势，达 90%左右，其中名词的比例最高，其次是动词，再其次是形容词，而儿童对表达语法意义的虚词的习得较晚。

大量的研究表明，儿童在各种词类的掌握上，最先掌握的是名词，而且是他们在日常生活中最先接触到的具体名词；其次是动词，1.5 岁以前就开始掌握，这也是比较具体的。稍为抽象的其他一些词类，儿童在 1.5 岁以后陆续掌握；形容词在 1 岁 6 个月到 1 岁 11 个月之间掌握；副词在 1 岁 10 个月以后；代词“你”、“我”在 1 岁 11 个月左右。更为抽象的数词和连接词，掌握得更晚些。

在 1.5 岁到 2 岁儿童所掌握的词汇中，名词和动词最多，占 70%，其他词类，包括形容词、副词、代词等，只占 30%。虽然不同的儿童最初说出的词可能有差别，但儿童早期学会的词大多意义非常具体，或是他们生活中经常见到的人和物的名称。如经常接触的人的称呼“妈妈”、“爸爸”、“姐姐”、“阿姨”等，身体器官名称“眼”、“耳”、“鼻”、“手”等，衣服名称“帽”、“鞋”、“袜”等，食物名称“饭”、“菜”、“糖”等。随着年龄的增长、儿童活动范围的扩

大以及思维能力的发展，儿童的名词量不断扩大。不仅具体名词的范围扩大，抽象名词也开始出现。

儿童早期的动词是用来描述正在进行的活动或刚进行过的活动，或用来表达自己的需要。儿童早期掌握的主要是动作动词，占绝对优势；其次是趋向动词、心理动词等，但比例较小。在其后的发展中，其他类的动词如能愿动词、判断动词、使令动词等逐步发展，但动作行为动词仍是儿童动词发展的主体。与成人不同的是，儿童经常用非动词来表达需要和描述活动。如儿童说"up"是表示他想让人把他举起来，说"嘟嘟"可以表示"开汽车"。

虽然有研究显示，儿童在 1.5 岁时已经能使用某些形容词，如"大"、"小"、"好"等，但没有一个形容词在这个年龄段的儿童中具有普遍性。"形容词的迅速发展，是儿童句子复杂化的一个标志，也是儿童对事物的性质认识迅速发展的一个标志"①。下表是朱曼殊等对不同年龄段儿童形容词使用数量的统计结果：

| 年龄(岁) | 使用词数 | 增长数 | 年龄(岁) | 使用词数 | 增长数 |
|---|---|---|---|---|---|
| 2 | 7 | | 4.5 | 108 | 48 |
| 2.5 | 21 | 14 | 5.5 | 159 | 51 |
| 3 | 41 | 20 | 6.5 | 206 | 47 |
| 3.5 | 60 | 19 | | | |

儿童最初掌握的形容词大都为描述物体特征的，儿童 2 岁时便已开始使用这类形容词，2 岁半开始使用描述动作、味觉、温觉和机体觉的形容词，3 岁时能使用描述人的外形特征、情感以及个性品质的形容词。4 岁以后是儿童形容词的快速发展期，4.5 岁时才开始使用描述事件情境的形容词。国外相关研究表明，儿童使用形容词的类型及其年龄段依次为：大小＞高矮/长短＞高低＞宽窄/厚薄＞深浅，这个观察结果与国内的研究基本一致：大小(2 岁)＞高矮/长短(3 岁)＞粗细(4.5 岁)＞低/厚薄/宽窄(5.5 岁)。儿童对形容词的习得和使用表现出从事物单一特征的描述到复杂特征的描述。如对人体外形特征的形容词"胖/瘦"3.5 岁就能使用，而对"老"、"年轻"等则分别要到 4.5 岁、5.5 岁时才能使用。

① 李宇明. 儿童语言的发展. 华中师范大学出版社，1995. 108

对于时间词的掌握，研究显示，3～4 岁时儿童能使用"先/后/同时"以及"正在"，4～5 岁掌握"昨天/今天/明天"的用法，"以前/以后"、"早晨/上午/中午/下午"、"已经"等 5 岁以后才能掌握，"去年/今年/明年"以及表示将来的"就要"等 6 岁以后才开始使用。3.5 岁的儿童会用"里"表示方位，4 岁左右学会使用"上/下"，5 岁以后基本掌握"后/前/外/中"等方位词的用法，"左/右"的习得较迟，要到 6 岁以后。

儿童代词的习得同样表现出普遍性，共同的特征是第一人称"我"的习得先于第二人称"你"，最后是第三人称"他"。儿童在指称方面最初使用直称，而使用人称代词需要比较高的抽象水平，儿童早期存在着"自我中心"的现象，因而对"我"的掌握先于"你"和"他"，"你"、"他"之间的指称关系比较抽象，因而不同民族儿童早期人称代词的习得过程中都不同程度地出现"你"、"我"倒置，第三人称"他"指代不清等。如：

成人：阿姨来了抱你吗？

儿童：抱你。（意为"抱我"）

成人：二姨打你了吗？

儿童：没打你，没打你。（意为"没打我"）

汉语儿童大约 3 岁时习得疑问代词"谁/什么/什么地方"的用法，4 岁时习得"什么时间/怎样"的用法，"为什么"要到 5 岁时才能掌握。值得注意的是，后两种疑问代词的习得与国外的相关研究不尽一致，国外相关研究显示，非汉族儿童习得"为什么"却先于"什么时候/怎样"。

汉语的数词与名词的组合必须借助于量词，汉族儿童大约 4 岁时已开始使用量词，但使用最多的是个体量词"个/只"，而且使用中呈现出"个/只"过度扩充的现象，如"一个汽车"等。4～5 岁的儿童个体量词的使用超过临时量词和集合量词，5 岁的儿童开始注意量词与名词之间的搭配，但也出现以动词作为临时量词的现象，如"一开汽车"等。6 岁儿童能根据事物的共同特征分别选择不同的量词，如把"车、飞机、轮船"等交通工具一律以"辆"来计量，或者把有轮子的物体都用"辆"来计量。量词有一定的抽象性，它们与名词或动词之间搭配的规律性较差，因而儿童掌握起来有一定的难度，直到 7 岁时儿童对量词掌握的程度仍然不是很高。

下表是许政援(1996)统计到的 1.5～3 岁儿童词汇中各类词的比例变化情况：

| 年龄(岁) | | 名 | 动 | 形 | 副 | 代 | 连 | 数 | 象声 | 语气 | 词尾 | 总计 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1.5～2 | 词数 | 366 | 299 | 62 | 88 | 41 | 6 | 11 | 9 | 6 | 62 | 950 |
| | % | 38.5 | 31.5 | 6.5 | 9.3 | 4.3 | 0.6 | 1.2 | 1.0 | 0.6 | 6.5 | 100 |
| 2～2.5 | 词数 | 287 | 354 | 55 | 102 | 145 | 7 | 14 | 4 | 27 | 70 | 1 065 |
| | % | 26.9 | 33.2 | 5.2 | 9.6 | 13.7 | 0.6 | 1.3 | 0.4 | 2.5 | 6.6 | 100 |
| 2.5～3 | 词数 | 208 | 237 | 62 | 96 | 151 | 12 | 5 | 4 | 33 | 52 | 860 |
| | % | 24.2 | 27.6 | 7.2 | 11.1 | 17.6 | 1.4 | 0.6 | 0.5 | 3.8 | 6 | 100 |

从上表可以看出，不同年龄段的儿童所习得的词汇初期以名词、动词、形容词等意义比较实在的词占绝对优势，但随着年龄的增长，比较抽象的词类特别是代词逐步增长；最抽象的数词和连词也稍有增长。这反映了儿童掌握语言从具体到抽象的普遍发展过程。

### 三、儿童词义的习得

词是语言中最小的能够独立运用的音义结合体，1～1.5 岁的儿童已能使用真正意义上的词来表达自己的愿望或情绪。这种运用不仅仅体现了儿童对词的形式（词的语音形式）的掌握，更为重要的是体现了儿童对不同的语音形式所代表的意义内容的理解。儿童对词义的理解是个逐步深化、不断完善的过程，初期表现出与成人对词义理解不完全相同的现象。

#### 1. 儿童对词义的过度概括现象

儿童通过自己所处的环境并在成人的语言教授下掌握了最早的一批词，这些词意义实在，所指对象也与儿童的生活关系密切。但儿童在对词所指对象及特征的把握上与成人的理解并不完全一致。当儿童将一个词从他听到的情境中推广或概括到类似的情境中去时，某些概括是适当的，如儿童将“狗”概括到狗的画或玩具狗等；但也有些概括是不适当的，如将“狗”概括到“猫”等。儿童理解词义的一种可能性是，抓住成人所说的词的意义的一部分，同时超越或扩充成人所指词义的范围。这种现象称为“过度概括”或“过度扩充”。

一般认为，“过度概括”大都发生在儿童语词所指的对象与该词的合适对象存在相似性的情况下。任何一个词都有许多语义特征，而儿童对于词的语义特征的掌握是逐步完成的。不同的儿童可以根据物体的不同特征来

扩充或概括一个词的意义，如有的儿童根据猫有4条腿、能行走等特征，将具有这些特征的牛、羊、狗等都叫做“猫”；也有的儿童根据猫的全身有毛这一特征，把羊、狗以及皮毛衣服都叫做“猫”。此类现象表明，儿童对某些词的意义的理解只是部分地与成人所理解的词义相同，而另一部分词义特征在成人看来是错误的。

儿童对词义理解的过度概括主要有两种方式：一是单一的过度概括，即以物体的一种至两种特征作为使用该词的标准，如有的儿童早期仅仅根据月亮是圆形的这一特征，把他所见到的所有圆形的东西都称为“月亮”，如皮球、柠檬乃至信封上的圆形邮戳和字母O等；二是混合的过度概括，即以和某个词原来所指对象在不同情境下所共有的某一特征为依据，过度扩充地使用这个词。如儿童用“踢”来指他自己踢一个静止的物体，其中至少包括3个特征：挥动肢体、突然碰到某个物体、物体被推动，后来只要包含这3个特征的现象都可以被儿童称为“踢”，如飞蛾拍动翅膀、自行车轮撞击一只球或者儿童自己抛掉某个东西等。

大量研究表明，儿童对词义理解中的过度概括依据的是物体的形状和功能，形状特征在最初可能是过度概括的主要原因，比如儿童更容易把“球”扩展到“苹果”，而较少把“汽车”扩展到“飞机”。物体的功能，特别是某些最普通的、在各种场合中都存在的功能，也常被儿童用作扩展的依据。

过度概括的发生是因为儿童只掌握了词的部分语义特征。因此当儿童说“狗”时，他所指的意义也许是“4条腿的动物”。以后随着生活经验的积累，他才掌握了把“狗”与其他4条腿的动物区别开来的其他语义特征。

2. 儿童对词义的概括不足现象

与“过度概括”相对，儿童对词义的掌握过程中，还存在着“概括不足”的现象，而且这种现象几乎和“过度概括”一样经常出现。所谓“概括不足”指的是儿童对词义的理解比成人对该词所指的范围要小，只是成人所指意义的一个子集。例如儿童把“车”仅限于指自己坐的车，“猫”只指家养的或邻居家中的特定的猫等等。有人认为，“这种意义范围的缩小，可能是儿童获得每个新词的初期表现”[①]。以后由于接触同类物体的机会增加，加上成人在交谈中所提供的一些非言语的或言语的线索，儿童逐渐能将其他的“车”和“猫”作类同理解，而且常常会出现“过度概括”的现象。

---

① 朱曼殊主编.心理语言学.华东师范大学出版社,1990.342

儿童在对词义习得过程中的过度概括与概括不足两种现象表明，儿童初期对于词和其所指对象之间的关系的认识是含糊的。当儿童开始产生比单词句长一些的语句时，儿童词义理解中的过度概括和概括不足的现象明显减少，这两方面的变化都与儿童认知发展的变化相吻合。儿童认知发展的变化表现在儿童从以综合的、普遍的态度对待物体和事件过渡到将物体或事件看成性质的集合，儿童开始将物体和活动分析成性质，并开始注意它们的个别性质。从这时开始，语言活动和环境不再以综合的、无区别的方式联系了。这一转变是儿童语言发展的重要里程碑。

*3. 儿童后期词义的发展*

儿童对词义的掌握随着他们语言系统的发展，必然逐步向目标语言靠拢，这是一个不断调整的过程。在这个过程中，儿童通过对词的语义特征增加或减少而改变词义理解中的过度概括或概括不足现象，这种调整与儿童认知水平的发展密切相关。

研究显示，儿童对单个词词义的发展表现在两个方面：一是不断充实该词的语义特征，从部分的、个别的语义特征向掌握较为全面的语义特征发展，以达到接近于成人的词义；二是从一个词的单义向多义发展，根据不同的情境使用该词的不同意义。如汉语中"前"、"后"两个词，可以表示空间关系，也可以表示时间关系。研究者观察到，儿童总是先理解这两个词的空间意义，然后才能理解它们的时间意义，其间相差竟达 1～2 年。因此随着时间的延续，儿童对单个词词义的理解是一个不断充实、加深的过程，最终与成人词义基本一致。

语言中词与词之间并非互不相关、各自独立的，儿童对词义的掌握不仅局限在对单个词意义的理解，还包括对语言中词与词之间意义关系的掌握。如在儿童词义充实、深化的过程中，掌握了"猫"与"狗"、"牛"、"羊"等词意义上的差别，另外逐渐习得了这些词与"鞋"、"帽"、"球"、"嘴"等词之间的差别。随着年龄的增长和认知发展水平的提高，儿童对词义的理解也从单个词的词义发展到对词与词之间意义关系的理解，将某些词所指称的对象归为一个类，从而与指称其他类对象的词区分开来，这又是更高层次上的意义理解与掌握。

## 四、儿童词汇、语义发展的特点

综合不同年龄段儿童词汇与语义发展情况，可以发现这种发展表现出

一些共同的特征和规律：

第一，儿童词汇的发展有先后的顺序，而且对情境的依赖性较强。不同民族儿童词汇发展的情况表明，虽然在某些词的发展上表现出一定的差异，如英语儿童习得"为什么"先于"什么时候"、"怎样"，而汉语儿童正好相反，但儿童对各种词的理解和产生有一个大致的顺序。

第二，儿童词汇的发展不匀称，而且有阶段性。儿童对词汇的理解和产生总的趋势是随着年龄的增长而增长，但并非总是呈上升趋势，而是表现出一定的阶段，有时快，有时慢，甚至有时会出现停滞的状况。而且在不同类词的习得上，发展也不平衡。

第三，儿童对词义的理解和掌握有一个逐渐精确化和概括化的过程。儿童对词义的掌握最初表现出两种表面上看似乎相反而实际相同的情况，或者是使某个词义过度扩充而发生词与词之间的混淆，如"前/后/左/右"不分；或者把词义过度缩小，使某个词只和某一个对象相联系，如仅仅将"狡猾"与"狐狸"联系起来等。这种现象表明儿童最初只是获得了词的部分语义特征，随着认知的发展，儿童会逐渐摆脱词与单一对象之间的联系，对词的全部语义特征逐步精确化和概括化，最终习得词的全部语义特征。

第四，儿童对词的理解和使用与儿童认知水平的发展，特别是概念的掌握密切相关。儿童对词的理解和使用表现出理解先于产生的语言发展的共同特征，不同语种儿童词汇的习得和使用具有一定的共性。虽然有一些发展顺序不一致的词，可能是由于语种不同，或者社会生活条件、文化背景的差异，或者是研究方法和材料的不同所致。

## 第四节　儿童语法的发展

### 一、句子的平均长度

句子是由词或短语构成的、具有交际功能的基本单位，每种语言都有一套将词组合成短语和句子的规则，即语法。语法规则的习得是儿童语言发展中的一个重要领域，通过语法规则的习得，儿童逐步掌握如何将词组合成更复杂的句子。儿童语法的发展通常从两个方面来评价和分析，一是句子的长度，二是句子结构的完整性和复杂性。

心理语言学界通常把儿童语言中出现的句子的平均长度(the mean length of uttrance, 简称 MLU)作为衡量儿童语言发展的一个重要指标，以此来反映儿童在不同阶段的语言能力的差别。所谓句子的平均长度指的是，在采集的儿童自发言语的样本中，对儿童的每一语句所包含的有意义单位的数目进行统计所得出的平均数。由于汉语缺少形态变化，汉字记录汉语不实行分词连写，因此我国研究者对儿童句于长度的研究有两种统计方法，一种是以字为单位，一种是以词为单位。吴天敏等和彭祖智等曾以字为单位来分析儿童句子长度的发展，结果显示，1 岁半至 2 岁阶段以 5 字以下句子为主；2 岁至 6 岁以 6～10 字句为最多。以字为单位计算句子长度，方法简便，易于操作。但汉族儿童早期语言中有许多重叠词或叠音词，如“爸爸”、“妈妈”、“帽帽”、“抱抱”等，这将会影响到对句子实际长度的评定。

儿童语言发展中句子平均长度的研究国外早在 20 世纪 30 年代就已出现，统计结果显示，不同年代儿童句子平均长度也有逐步增加的趋势，有人将此归结为“由于促进语言发展的条件增加了，并且容许儿童和成年人交谈的机会也更多了，所以语言也就学得更好”[①]。国内学者朱曼殊等也曾以词为单位统计过汉族儿童的平均句长，结果如下：

| | 2 岁 | 2.5 岁 | 3 岁 | 3.5 岁 | 4 岁 | 5 岁 | 6 岁 |
|---|---|---|---|---|---|---|---|
| 朱曼殊等 | 2.91 | 3.76 | 4.61 | 5.22 | 5.77 | 7.87 | 8.39 |

对于句子长度对评价儿童语言发展水平的作用，朱曼殊等(1990)、李宇明(1995)认为，这虽然是一种通用的评定儿童早期语言发展的指标，但只是一种外在指标和次要的指标，因为它只能表明句子中所含的字、词在数量上的发展，而不能反映儿童语言发展或语法发展的质的变化，不能真实地反映句子的结构性质和复杂程度。儿童进入完整句阶段后，同样长度的句子可能具有完全不同的结构和复杂程度。另外，这种测定貌似客观，但词、短语、句子等语法单位之间存在划界的问题，测定过程中会因研究者的不同而有所差别。因此，句子结构的完整性和复杂程度以及类型差别才是衡量语法发展的更为重要的指标。为此，吴天敏、彭祖智、朱曼殊等和“中国幼儿口头语言研究协作组”分别采用纵向研究和横向研究的方法对儿童语言发展中

① 卡罗尔(J. B. Carrol). 儿童语言的发展. 曾越麟译. 语言学动态. 1979(4)

句法结构的发展作了研究，并按照儿童语句的完整性和复杂程度，将句法结构分为不完整句、完整单句和复合句 3 个阶段。认为 3 个阶段之间的界限具有过渡性，有相互重叠的部分。

## 二、不完整句阶段

儿童语言发展中的不完整句阶段又可以分为单词句（独词句）阶段、双词语句阶段和电报句阶段。

1. 单词句阶段

1 岁左右的儿童对一些经常接触的人或物已经能正确地称呼，如看到父母时能分别叫出“爸爸”、“妈妈”；要大人抱时，会伸出两臂叫“抱抱”。随后不久儿童会用单词来表达自己的愿望或要求，儿童语言的发展进入了单词句阶段。

单词句阶段儿童所使用的单词已不限于词的指称功能，这一阶段的儿童用单词去表达自己的需要、愿望或情绪，或用单词来描述周围的情境或事件。这些具有交流功能的发音，虽然形式上只是一个单词，但所表达的意思往往相当于成人所说的一句话，故称单词句。

单词句有几个特点：一是和动作紧密结合。当儿童用单词句表达某个意思时，常伴随着动作和表情。二是意义不明确，语音不清晰，成人必须根据当时的情境和语调等线索才能推测出意思。三是词性不确定，虽然儿童最先学到的是名词和动词，但在使用时并没有明确的词性分别，一个名词可以用来指物，也可用来指动作行为事件等。如“帽帽”可以指帽子，也可以指戴帽子；“嘟嘟”既可以指汽车，也可以指开汽车等。

不同的学者对儿童出现的单词句的性质的认识有分歧，有人认为这个时期儿童所说出的单词仅仅是一个单词而已，并不具有句法基础；也有人认为在单词句阶段，儿童已具有了简单的句法关系和句法范畴的知识，不能说出完整的句子并非由于不知道句法，而可能是没有掌握足够的词和受到其他方面的限制。朱曼殊等根据对单词句儿童的追踪观察发现，儿童的第一批单词和单词句之间在丰富性和语用功能上存在着区别，后者不仅仅是利用单词去给事物命名，而且具有明显的表达意图，虽然在形式上与成人的句子存在着较大的差别，但在功能上类似于完整的句子。

由于此阶段儿童所发出的表达意义的单位有时难以界定是不是一个词，李宇明提出用“语元”概念代替独词句阶段的“词”。他通过对一

个女孩D的个案材料的调查,将这一阶段的语元分为4种类型:指称、行为(包括动作)、否定、语气。结果显示,其中指称性语元最多,包括指人的、指物的和指称动物的;其次是行为语元和否定语元,否定性语元包括“不”、“不喝”、“我不”、“不敢”、“不要”、“不穿”、“没有”等,除了“没有”用于对存现的否定之外,其他都用于对动作行为的否定,多用来否定儿童的主观意愿;语气语元数量最少,主要用于表示应答和惊疑,有“呀”、“唉”、“嗯”3个。

2. 双词句和电报句阶段

经过单词句阶段的准备,儿童到1.5岁左右开始说出由两三个词组合起来的语句,如“妈妈鞋”、“宝宝帽帽”等等,儿童语言进入了双词句阶段。这一阶段,儿童的语言是新颖的。例如儿童没有听到过“妈妈鞋”或“宝宝帽帽”这一类的话语,他们听到的可能是“妈妈给宝宝穿鞋”、“这是妈妈的鞋”,“妈妈给宝宝戴帽帽”、“这是宝宝的帽帽”之类。因此,儿童的这些语言是在没有范例的情况下自己创造出来的。这一阶段是所有儿童都要经历的一个由单词句阶段向完整句发展的重要过渡阶段,这个阶段儿童所使用的词仍以名词和动词等实词为主,连词、介词、助词等功能词很少出现。在表意上,双词句以及经过有限扩展的多词句虽较单词句明确,但其形式是断续的,结构不完整,类似于成人的电报文本,又被称为电报句。

双词句或电报句和单词句的本质差别不在于句子形式的长短,而在于儿童的双词句中包含着组合的因素,体现出一定的组合功能。但对于双词句是否具有内在的组词成句的规则,学者们提出了轴心语法说和语义关系说等理论。

轴心语法说认为,儿童的双词句中包括两类词,一类词数量不多,范围较小,比较稳定,增长速度较慢,但使用频率很高,而且在双词句中出现的位置相对固定,可以称为“轴心词”;而另一类词的数量较多,范围较大,变化较多,增长较快,但使用频率较低,在双词句中出现的位置比较自由,可以称为“开放词”。双词句以轴心词为中心组织起来,构成“轴心词+开放词”、“开放词+轴心词”等词与词的组合。前者如英语中的“Allgone Daddy”,“Allgone milk”,汉语中的“妈妈抱”、“妈妈鞋”、“妈妈糖”等,其中“Allgone”,“妈妈”是轴心词。后者如英语中的“push it”,“close it”,汉语中的“吃糖”、“吃饼饼”等,其中的“it”,“吃”是轴心

词,其他的词是开放词。

轴心语法说注意到了这一时期儿童语言组合方面的一些特点,以轴心词为中心的双词句是儿童双词句构成的重要模式,甚至有人认为轴心语法是儿童语言发展中具有普遍性的第一步,是儿童语法向成人语法发展的基础。但随着研究工作的深入,这种学说也逐渐显示出它的局限性:

1) 如何区分轴心词与开放词?事实上,这两者之间的界限往往很难划分。

2) 儿童双词句中轴心结构出现的频率和比率并不高,并不具有普遍性。

3) 儿童双词句中的轴心词分属不同的词类,以英语为母语的儿童语言发展状况表明,充当轴心词的有指示代词(that)、人称代词(it)、数量词(more)、介词(off)、动词(want、to)等。而不同性质的轴心词与开放词的组合,其语法关系是不同的。轴心语法说不能说明这些词与开放词组合的语法关系,因而也不能从根本上说明儿童是如何从双词句过渡到成人语言的语法结构的。

4) 轴心语法说只强调语法规则而忽视了句子的意义,不能说明儿童双词句中轴心词与开放词之间的语义关系,特别是同一个双词句在不同的语言环境中可以表达不同的语义关系这一现象。如"mommy sock"这一双词句在不同的语言环境中可能表达施事-受事的关系,相当于"妈妈帮助儿童穿袜子";也可以表达领属关系,相当于"妈妈的袜子"。轴心语法说并不能反映双词句中的这些意义关系。

针对轴心语法说在儿童双词句性质解释方面的某些不足,有学者提出了语义关系说。这种学说认为,儿童在双词句中所表达的,是儿童在感知运动阶段以对事物间关系的认识为基础的语义关系。儿童在双词句阶段所表达和了解的不仅要比两个单词的意义多些,而且还了解两个词之间的意义关系,并用一定的词序来表达这种关系。如说英语的儿童在表达施事和受事的关系时,总是把施事放在受事的前面,在表达领属关系时,总是把领有者放在前面,领有物放在后面。因此双词句中单词的出现顺序多半是根据各个词在话语中所起的作用来安排的,并非像轴心语法说所说的那样一个一个地去学习各个词的固定位置。

布朗发现儿童在双词句中所表达的意义具有高度的一致性,并总结了11种表达意义的形式,按性质可分为指称形式和关系形式两大类:

| 形　式 | 表达意义 | 例　句 |
| --- | --- | --- |
| 指称形式 | 叫出名字 | 这狗狗 |
| | 反复 | 还要糖糖 |
| | 消失 | 球球没了 |
| 关系形式 | 施事和动作 | 妈妈抱 |
| | 动作和对象 | 开嘟嘟 |
| | 施事和对象 | 妈妈(穿)鞋鞋 |
| | 动作和位置 | 坐凳凳 |
| | 物体和位置 | 饭饭碗 |
| | 物主和所属物 | 妹妹娃娃 |
| | 物体和属性 | 好看衣服 |
| | 指示词和物体 | 这个娃娃 |

语义关系说还提出了在双词句基础上向完整句发展的两种可能方式：一是将两种语义关系连接在一起而省略其中共有的一个词。如：

(施事＋动作)＋(动作＋对象) ——→ 施事＋动作＋对象
妈妈　抱　　　抱　　弟弟　　　妈妈　抱　弟弟

二是将关系中的一个元素扩展成关系本身。如：

对象——→ 物主＋所有物
帽帽　　宝宝　帽帽
动作＋对象——→ 动作＋物主＋所有物
拿　帽帽　　拿　宝宝　帽帽

语义关系说考虑了儿童话语的意义，提供了早期儿童语言发展普遍性的证据，并能对早期儿童语言发展的连续性作出比较合理的说明。

双词句和电报句是儿童语法发展由不完整句到完整句的过渡阶段，通过这一阶段，儿童开始建立句子的基本模型，语言中出现了一些语法类化现象，并且句子所表达的意义呈现出复杂化和明显化的迹象。

### 三、完整句阶段

经过不完整句阶段的准备和调整，儿童语言逐渐向成人语言靠拢，进入

了完整句阶段。研究显示,在不完整句阶段,就表现形式和主要功能而言,以不同语言为母语的儿童的语言发展表现出较强的一致性。但在完整句阶段,则体现出不同语言之间的差别和特殊性。如形态变化比较丰富的语言,儿童语言的发展表现出在语法方面出现两种变化:一是句子的长度和结构的完整性与复杂性增加,开始按照一些基本的语义关系将单词组成完整句;另一方面,在单句中开始出现一些词形变化。汉语是典型的分析性语言,缺乏形态变化,汉语儿童经过不完整句阶段后,其语法的发展过程与英语儿童存在着明显的差别。

1. 单句的发展

单句分为简单单句和复杂单句两种。句子根据语气可分为陈述句、疑问句、祈使句和感叹句 4 类,儿童最初产生的大多为陈述句,其他句类的比例很小。

(1) 简单单句

1.5 岁到 2 岁左右的儿童在说出电报句的同时,开始说出结构完整但无修饰语的简单单句。如"娃娃觉觉。""宝宝吃糖糖。"

汉族儿童简单单句阶段的特点和明显标记是:儿童的单句从无修饰语到各种修饰语的出现。乐守红[①]认为,2 岁以后,儿童语言中有修饰语的语句表现出随年龄增长而逐渐增多的趋势。从修饰语的类型看,定语较多,状语次之,补语最少;充当定语、状语、补语的有单词,也有短语;有单层的,也有多层的,而且多层定语的出现早于多层状语。如:

爸爸你看,那是哪吒的爸爸。(2.0)
我有一个小汽车,红色的小汽车。(3.0)
这个枪你要不要?我的枪太大了。(2.0)
爸爸,你今天跟我一起上学去好不好?(3.0)
这树叶脏死了。(2.0)
我的肚子挺得老高老高的。(2.0)

有研究表明,到了 3 岁左右,儿童开始使用较复杂的名词性结构"的"字句和"把"字句,如"这是我玩的玩具"、"我把积木放在盒子里";与此同时,还出现了较复杂的时间及地点状语,各种语气词也开始出现,如"你坐到沙发

① 乐守红.2 岁至 3 岁儿童语言发展研究.南京师范大学硕士学位论文.2003

上去吧"、"这有什么了不起啊"。研究者还发现，汉语儿童3.5岁时，在单句中使用复杂修饰语的句数和修饰语的种类增长速度最快，约为3岁时的两倍，以后直到6岁逐年增长，但增速不明显。也有研究者据此认为，3.5岁可能是汉族儿童简单单句发展的关键期。

(2) 复杂单句

复杂单句的特点是突破了简单单句的"主-谓"、"主-动-宾"、"动-宾"等无修饰成分或只有简单修饰成分的模式，出现了复杂短语充当谓语或其他句法成分的结构。朱曼殊等(1990)研究发现，在2岁到6岁的儿童语言中出现了3类复杂单句：一是由几个动词结构连用的连动句，如"小朋友看见了就去告诉老师"；二是由一个动宾结构和主谓结构套叠的兼语句，如"老师教我们做游戏"；三是句子中的主语或宾语中又包含主谓结构，如"两个小朋友在一起玩就好了"。

通过研究，朱曼殊等归纳了儿童各种单句发展的一般顺序：不完整句；具有主-谓、谓-宾、主-谓-宾、主-谓-补等结构的无修饰语单句；简单修饰语单句，主-谓-双宾语句，简单连动句；复杂修饰语句，复杂连动句，兼语句；主语或宾语含有主-谓结构的句子。

**2. 复合句的发展**

复合句在2.5岁儿童的语言中就已有少量出现，到了5岁就已发展得较为完善了。复合句阶段的特点是儿童可以将两个单句根据它们之间的逻辑关系排列成句，但是结构松散，缺少关联词语，一般是无标记的复合句，以联合复句为主，偏正复句所占比例较小。

联合复句中出现最多的是并列复句，即把两件并列的事加以陈述，如：

我妈妈小，我爸爸大。(2.5)

我爷爷也没有票，我奶奶也没有票。(2.0)

其次是连贯句，按事情的经过描述所发生的情况。如：

把大石头搬起来，把大灰狼打死。(2.0)

西瓜杀开了，猪八戒吃了。(3.0)

偏正复句中出现较多的是因果关系的复句。如：

我家没有饭，给你吃辣椒。(2.0)

这枪没子弹了，修好了再给你。(3.0)

转折复句和条件复句也有出现，但数量较少。如：

我叫他不要推，他非要推。(3.0)

我给他枪，他要我就给。(2.0)

复合句的组合有意合法和形合法两种方式，儿童复合句的发展初期以意合法为主，通过关联词语将几个分句组合起来的形合法的发展稍晚。关联词语运用初期，儿童使用得较多的是"还"、"也"、"又"、"就"等具有关联功能的副词，后来才出现"那么"、"只好"等，到5～6岁时使用的关联词语频率有所增加，出现如"因为"、"为了"、"如果"、"可是"等说明因果、转折、条件、假设等关系的连词以及前后呼应的成对使用的关联词语，如"一边……一边"、"没有……只有"、"要不……就"等。另外，儿童对关联词语的使用经常有误，有时是使用不当。如：

我家里昨天一共来了好几个客人，反正都是认识的人。

你不把积木整理好，要不然我去告诉老师。

有时则是形式和内容不一致，如：

告诉他要下雨了，他就回家去睡觉。

他还会做算术，还做那个加法、减法。

这些情况说明，关联词语的获得过程相当复杂，它和儿童逻辑思维的发展水平密切相关。

3. 句式的发展

单句和复合句是从句子结构角度划分出来的类，通常称为句型。每种语言中还有一些有特殊的构成标志或构成模式的句子，通常称为句式。在儿童句式发展研究方面，近一二十年取得了丰硕的成果，一方面丰富了人们对儿童语言习得规律的认识，另一方面，不少学者将通过对儿童句式发展情况的研究，来验证、补充、修改已有的语法研究的观点。以下仅以汉语儿童"把"字句和被动句的发展为例，说明儿童语言研究对汉语语法学理论的贡献。

(1) 儿童"把"字句的发展

"把"字句是汉语中很有特色的一种句式，研究成果也较多。对于这种句式的性质，传统上解释为通过介词"把"将动词的宾语提前，整个句子带有处置意义。但语言事实中有些"把"字句并不能还原为一般的"主-动-宾"

句，也有些"把"字句用处置说不能做出令人信服的解释。通过对儿童"把"字句习得过程的分析，可以为这种句式提供更为完满的解释。

关于汉语"把"字句的构成，通常认为必须具备以下几个条件：① "把"字句中的动词不能是光杆形式，或者前有状语，或者后有补语，至少要有动态助词；② 动词带有处置性；③ "把"介引的对象通常是有定的；④ 否定副词、能愿动词等只能出现于"把"前，不能出现于"把"字介词结构与动词之间。这几个方面的条件决定了儿童对"把"字句的习得只能在双词句以后的语言发展中才有可能出现。据李向农、周国光①等的研究，"把"字句在 2 岁儿童中已经出现，4～5 岁基本发展成熟。这一阶段也是儿童完整句发展的高峰期，句法结构、语义关系等的复杂化、母语化为"把"字结构的发展完善创造了必要的条件。

周国光、王葆华(2001)将"把"字句细分为 9 大类，并分别考察了儿童对不同类型的"把"字句的习得过程。这些类型是(每种类型最后括号内的数字表示该句式始见的年龄段)：

1a. 把＋名词＋动词＋趋向动词(2.0)

1b. 把＋名词＋动词＋动词/形容词(2.0)

2. 把＋名词＋在/到＋处所名词(2.0)

3a. 把＋名词＋副词＋动词(2.5)

3b. 把＋名词＋一＋动词(3.0)

3c. 把＋名词＋给＋动词(3.0)

3d. 把＋名词＋介词词组＋动词(3.5)

4a. 把＋名词＋动词＋了(3.0)

4b. 把＋名词＋动词＋着(3.0)

5a. 把＋名词 1＋动词＋名词 2(名词 2 属于名词 1)(2.5)

5b. 把＋名词 1＋动词＋名词 2(名词 1 表物，名词 2 表人)(3.0)

5c. 把＋名词 1＋动词＋名词 2(名词 1 表处所，名词 2 表动作的结果)(3.0)

5d. 把＋名词 1＋动词＋名词 2(名词 2 表示名词 1 的一部分)(4.5)

6. 把＋名词＋动词重叠(2.5)

---

① 李向农. 2～5 岁儿童运用"把"字句情况的初步考察. 语言研究. 1990,(4)；周国光. 汉语句法结构习得研究. 安徽大学出版社，1997

7. 把＋名词＋动词＋得＋情态补语(4.0)

8. 把＋名词＋动词＋动量(4.0)

9. 把＋名词 1＋当/V 成＋名词 2(4.5)

通过调查与分析,发现从 2 岁开始儿童"把"字句的使用随着年龄的增长,"格式由简而繁,构成成分由少而多,句子长度由短而长",而且 1a,1b 及 2 三种类型一直是使用频率最高的基本结构,贯穿于整个发展过程。到 4.5 岁时,除了成人语言中的"把＋名词＋动词＋了＋动词"和"把＋名词＋动词＋时量"两种格式未出现外,已与成人的"把"字句状况基本接近。

(2) 儿童被动句的发展

儿童被动句的发展需要一定的心理基础和语言基础,周国光(1997)认为,儿童被动句始见于 2 岁年龄段,这种句式的发展体现在结构、语义、功能 3 个方面。

结构方面包括结构成分、结构层次、结构关系等。早期儿童使用一些结构比较简单的被动句,儿童以这些结构为基础,运用扩展、替换、联结等手段实现了被动句的发展。如在基础结构中的受事主语、动词前添加附加成分、用一定的成分去替换基础结构中的某一部分、在基础结构的谓语部分接续上新的成分等。如:

| | |
|---|---|
| 石头甩进去 | 石头甩洞里 |
| 小鸭放下 | 小鸭放在这儿 |
| 这手搞破了 | 这手搞破了一块 |
| 他给坏蛋打了 | 他给坏蛋打伤了一条腿 |
| 大蛋给我＋我吃大蛋 | 大蛋给我(我)吃(大蛋) |
| 书买来＋看看书 | 书买来看看(书) |
| 钱装进去＋钱倒不下来 | 钱装进去(钱)就倒不下来。 |

语义方面包括语义成分、语义内容、语义关系等。语义的发展同结构的发展是一致的。通过扩展、替换、联结,语义成分增加了,语义内容丰富了,语义关系也复杂了。在儿童被动句的基础结构中,最基本的语义关系是"受事-动作"。儿童在习得这一主谓结构以后,就利用它去包摄更多的语义关系和语义内容。如:球拿到了、大马骑过了、飞机叠好了、故事讲完了、颜色涂好了、顺序摆乱了等。

通过对 2 岁年龄段儿童语言中被动句及其与主动句之间的转换关系的

考察与分析，发现两者之间同样存在转换不对应性。这种不对应现象表现在3个方面：

一是并非每一类被动句都可以转换为相应的主动句。如：

球搞到这个里面去了　？搞球到这个里面去了

帽帽放叔叔口袋里　？放叔叔口袋里帽帽

二是在每一具体类型里，也因具体语句的不同而使转换出现不对应现象，由某些被动句转换成的主动句语感别扭。如：

这个捆起来啦　？捆起来这个啦

大灰狼也搞死　？也搞死大灰狼

手绢藏起来了　？藏起来手绢了

三是被动句和相对应的主动句出现的时间不对应。如“名受＋动1＋名兼＋动2”类被动句可以转换为相应的主动句：大蛋给我吃—给我大蛋吃，但是与之相对应的主动句句型半年以后(2.5岁年龄段)才出现。如“我妈妈给我鱼吃/给你一个小叶子玩玩”。

(3) 从儿童对“把”字句、被动句的习得看主动句与被动句的转换

“把”字句和被动句出现于相同的年龄段——2岁时开始使用，通过儿童“把”字句和被动句习得过程的分析，不仅初步勾勒了这两种句式的发展过程，更为重要的是对长期流行于汉语语法学界的主动句与被动句、“把”字句与“主-动-宾”结构之间存在转换关系的说法提出了质疑。通过对2岁年龄段儿童语言中的“把”字句、被动句和主-动-宾句的比较，周国光(1997)认为：就3种句式中的受事成分的性质而言，“把”字的宾语、受事主语句的主语基本上是有定的，而“主-动-宾”句中的宾语既有有定的，也有无定的；就句中动词的性质而言，“主-动-宾”句中的动词几乎都是单纯的，即光杆动词；同期出现的“把”字句、被动句中的动词都不是单纯形式，而要求是动结式、动趋式等复合形式，或要求动词后附加“了”、“过”等动态助词，或要求主谓短语作谓语。这种动词形式上的差异极大地限制了“主-动-宾”句式向“把”字句、被动句的转换。主动句和被动句在句法构成上的区别也决定了两者在转换上的不对应性，而两者在句法构成上所要求的条件不同也决定了习得两者的语言基础不同。

通过儿童主动句与被动句习得情况的调查与分析，周国光(1997)对长期以来有关汉语主动句、“把”字句与被动句之间关系的论述提出质疑，认为

"汉语被动句同'把'字句关系密切,被动句在主动-被动关系上不同主动句相对应,因而我们不同意主动句转换习得被动句的假设","恰恰相反,我们倒倾向于认为'把'字句是通过受事主语句增加受事标记而习得。因为受事主语句的结构要相对简单,习得难度也较小,所以也较早习得"。

## 参考书目


1. 桂诗春. 心理语言学. 上海外语教育出版社,1985

2. 靳洪刚. 语言获得理论研究. 中国社会科学出版社,1997

3. 李宇明. 儿童语言的发展. 华中师范大学出版社,1995

4. 李宇明. 儿童语言发展的连续性及顺序性. 汉语学习. 1994,(5)

5. 彭聃龄. 语言心理学. 北京师范大学出版社,1991

6. 许政援. 3 岁前儿童语言发展的研究和有关的理论问题. 心理发展与教育. 1996,(3)

7. 许政援. 对儿童语言获得的几点看法. 心理发展与教育. 1994,(3)

8. 许政援等. 儿童发展心理学. 吉林教育出版社,1987

9. 周国光、王葆华. 儿童句式发展研究和语言习得理论. 北京语言文化大学出版社,2001

10. 朱曼殊主编. 心理语言学. 华东师范大学出版社,1990

11. 周国光. 汉语句法结构习得研究. 安徽大学出版社,1997

# 第八章　地名学和人名学

## 第一节　概　　述

### 一、地名和地名学

1. *地名*

(1) 地名和地名系统

地名是专门指代大大小小的地域的语言符号。地名指称的对象不是地理实体，而是个体地域。组成地理环境的有形物体都是地理实体，比如高原、平原、盆地、河流、湖泊、海洋等是自然地理实体，城市、乡村、厂矿、铁路等是人文地理实体。地理实体固然处在一定的区域，但是，地理实体所处的位置仅仅是地理实体的一个方面，地形、地势、水文、植被、气候等都是自然地理实体研究的对象，而且是更为重要的研究对象。与此不同，地名所研究的主要对象恰恰是各种地理实体所处的地域。比如"昆仑山"这个地名，指称的是昆仑山这个地理实体所处的区域，并没有指称昆仑山的地质、地貌、走向、植被、气候等方面特征的功能。所以，地名的本质功能在于指代方位。

从地名的存在形式上来看，地名在产生之初，只有口语的表达形式，文字产生之后，有了文字形式。地名产生于社会生活，应用于社会生活，地名要便于人们称说和记录，所以，地名的读音、字形和意义也就成了地名必不可少的组成部分。总而言之，位、音、形、义是地名的构成要素，四者缺一不可。

地名指代的是一个一个具体的地域，数量庞大，但是，地名并不是一盘散沙，而是一个包含不同层次、不同类别的方位指称系统，这个系统就是地名系统。地名系统实质上是一系列大大小小的地域名称组成的集合。比如

从地名所指称的对象来看，地名可以分为自然地理地名系统和人文地理地名系统，自然地理地名系统根据不同的地理实体又可以分为表示陆地的地名系统、表示水体的地名系统；人文地理地名系统又可以分为行政区划地名系统、聚落地名系统、建筑地名系统和文教设施地名系统。陆地、水体、行政区划、聚落、建筑、文教设施等地名系统还可以划分成更小的系统。再比如从地名的语言形式上来看，地名系统可以分为结构系统、语源系统和语用系统。地名的结构系统又分为原生地名系统和相关地名系统，前者又可以细分为单纯地名系统、复合地名系统和词组地名系统，后者可以进一步分为转类地名、移用地名、派生地名和拼合地名。地名的语源系统又包括汉语地名和外语地名，汉语地名又有共同语地名和方言地名之分，外语地名又有民族语地名和外来语地名的差别。地名的语用系统包括现代地名、历史地名和虚拟地名。

(2) 地名的属性

1) 地名的语言属性。地名是人们共同约定的一种语音符号，属于语言范畴，语言性是它的首要特征。地名的语言性表现在：第一，语言文字是地名存在的物质基础，地名必须借助于语言和文字来表示。地名和语言中的其他词语一样，具备音、形、义三要素。第二，语言是有民族性的，因而地名作为一种语言符号，地名的语言性还表现在地名的民族性上。地名一般总是由当地的居民以自己的语言命名的，所以通过地名词语特点的分析，可以帮助鉴别地名的语别，而语别的不同一般又表示不同民族的分布范围，所以，地名的语言特征可以反映领土的归属。比如钓鱼岛、黄尾屿、赤尾屿等岛屿，早在 1534 年就见于我国史册，这些名称早就出现在日本所称的尖阁列岛之前，历来是我国的领土。

2) 地名的地理属性。地名是地域的名称，用来指称个体地理实体所处的区域，而地名所指的地方也是地理学研究的内容之一，因而地名又有地理属性。地名一般由通名和专名组成，通名指出地名所属的地理类别，比如“太行山”、“峨眉山”中的“山”是通名，作用在于指明类别；专名用来区别同类地理实体的不同个体，“太行”、“峨眉”是专名，使这两座不同的山得以区别。一般来说，通名用来定类，专名用来定位。有的地名反映了地理事物的特点，比如黄河、赤水，反映了河水的颜色，葫芦岛、月牙泉反映了地形特点，铁岭、铜陵反映了矿藏，南京、北京反映了相对位置。还有的反映了区域地理特征，比如反映黄土高原地貌的通名有“塬”、“峁”、“梁”，反映长江三角洲

河网密布地理特征的"浜"、"汇"、"浦"等等。

3) 地名的社会属性。名称是为了指称和交际，地名也不例外。人类居住和社会活动离不开场地，为了标志和识别各种地理事物，便于社会成员共同使用，就产生了地名，地名是全社会共同创造的。只有当大多数社会成员接受、认可了某一个名称，这个名称才能作为地名起到指称和交流的作用；而地名一旦被社会成员接受，就具有了相当的稳定性。据统计，我国历史上每一个朝代的地名变更率一般为百分之一到百分之二，最多不超过百分之三。所以，地名不能随便改动，否则会给人们在行政、邮电、交通、治安管理等方面造成很大的不便。但也正因为地名有社会属性，所以地名也不是一成不变的。随着社会历史的发展，有些地名也会发生变化。比如北京公元前 1000 年就称为蓟，后为燕国的统治中心。秦朝置广阳郡，汉晋称为燕国，唐代称为幽州，辽会同元年称为幽都府，建号南京，又称燕京。金贞元元年，定都于此，改名中都。元至元九年，改称大都。明太祖改为北平，明永乐元年改为顺天府，建北京，19 年后迁都于此，改称京师。辛亥革命后改了几次名，建国后定称北京，沿用至今。

地名的社会性还要求地名简洁，易读、易写、易记，以便交际。内蒙古的一些荒僻地区的地名，如果用汉字译写，显得很累赘，不利于交际。比如"楚鲁特达巴格音珠恩高吉格尔"(Qulut Dabagin Jun Gojgor，意思是石头山岭左边的尖山)。

(3) 地名的功能

人类的活动离不开地名，地名在人类生活的方方面面都发挥着作用。首先，地名是社会活动的基础。在人类语言产生的同时，地名也就产生了。随着人类活动范围的扩大，地名也越来越多，越来越复杂。正是因为有了地名指称不同的活动区域，才使人类的活动空间得以明确区分，生产得以展开，商品得以流通，可以说，地名是人类社会的基础。试想，如果没有地名，异地的人们怎么联络？没有地名，贸易怎么进行？没有地名，人们怎么出行？

其次，地名是国家行政管理的工具。世界各国都是通过划分行政区域、设立各级政府来管理国家的。行政区划都是通过相应的地名表现出来的，一个表示行政区划的地名，就指称一个确定的地域。如果没有行政区划地名，国家是难以实现对全国领土实施有效的管理的。

第三，地名是记录历史发展的化石。地名是人类活动的产物，在漫长的

人类发展过程中，有的地名忠实地记录着人类社会历史和自然演变的痕迹，例如元末明初，由于战乱灾荒，河北平原人口锐减，土地荒芜。为了充实此地的人口，尽快恢复农业生产，明王朝从山西实施了大规模的移民。据文献记载，明洪武二十二年、三十五年及永乐二年、三年，明王朝从山西大量移民到北京地区。现在北京的大兴县和顺义县的一些地名反映了这一历史事实，比如大兴县东部凤河沿岸有许多村子的名称是以山西的州县命名的，如长子营、沁水营、河津营、潞城营、蒲城营、霍州营、赵县营、黎城营、解州营、大同营、山西营等，顺义县西北部有夏县营、绛州营等地名。这些地名有力地证明了现在北京市的大兴县东部和顺义县西北部是明王朝安置山西移民的主要地区。再比如南京市建于明朝的城墙原来有内城和外城，现在外城荡然无存。可是在距离市区很远的仙鹤门、麒麟门、上方门、尧化门等地名证明了当时南京外城的广阔范围。地名还是自然环境演变的活化石。有的地名记录了河流的改道。比如湖北省的汉阳县，根据"地在水之北曰阳，在水之南曰阴"的命名原则，汉阳县应该在汉水的北面，可是现在汉阳却在汉水的南面，这是为什么呢？在明成化元年（公元 1465 年）以前，汉水入江地点在今汉阳县的南面，成化初年，汉水改道，在汉阳北面流入长江。所以，汉阳县这个地名对汉水改道提供了确凿的证据。再比如内蒙古的大青山。大青山名为"大青"，说明这座山应该郁郁葱葱，但实际上大青山上树木很少。但是在古代，这里却是草木繁茂的地方。据《汉书·匈奴传》记载，匈奴人曾把这里当作自己的苑囿。蒙古人来到阴山的时候，这里仍然草木繁盛，"呼和浩特"蒙语是青色的城，"包头"蒙语是有鹿的地方，这两个地名告诉人们，在古代，内蒙的大青山还是一个生态环境美好的地方。

第四，地名还能反映丰富多彩的文化内涵。据李如龙统计，安、宁、兴、昌都是汉语地名很常用的字，它们反映了中国历史文化传统中对大一统的向往；龙是中华民族的图腾物，因而龙在地名中出现的频次也非常高；道教是中国土生土长的宗教，道教崇拜仙人，因神仙得名的山岭洞府、河湖津渡就有许多，比如九仙山、仙都山、仙霞岭、仙人洞、仙溪镇等等，反映了道家的宗教思想。

2. 地名学

地名学是综合研究地名的起源、含义、词语特征、演变和分布规律及其应用的一门学科。地名学以地名系统为研究对象，而地名是人类历史的产物，是一种独特的社会现象，因而首先属于社会科学范畴。

地名学和语言学关系密切。通过分析地名词语在音、形、义方面的特征，可以发现地名中包含着许多语言学知识。许多地名保存着古汉语的词义。比如“江”在古汉语中是一个专有名词，指长江。这个古义还保存在许多地名中，比如“江阴、江陵、江宁”，都在长江两岸。再比如“汤”在古汉语中是“热水”的意思，普通话中已经没有这个意思了，但是在一些地名中还保存着这个古义，比如南京附近的汤山、汤泉，北京的小汤山，都以温泉而著名，其中的“汤”显然是热水的意思。许多古音也保存在地名中，比如江西的铅(yan)山、广东的番(pan)禺都是古音的保留。同样，地名学研究也需要语言学。研究地名首先要从调查地名的形、音、义入手，不掌握语言调查、方言调查的方法，就无法下手。分析地名词语的结构，没有语法学理论的指导，很难做到科学化；研究地名的起源，没有古今语言知识和亲属语言知识，更是寸步难行，所以地名学需要有语言学的支持。

地名学和地理学的关系也非常密切。地名往往反映了一个地方的地理类型和特点，通过地名研究，不仅可以了解当代的地理状况，而且有助于了解古代的地理面貌，了解地理特征的发展演变。同样，如果不了解一个地方的地形、地貌、人文、经济等地理特征，也很难透彻了解一个地名的意义。所以，地名学的研究也离不开地理学。

地名是历史的产物，地名中包含着许多历史知识。江苏省宿迁县本名宿豫县，到了唐代，因为要避唐代宗李豫的讳，所以改名宿迁县；宜兴县原名为义兴县，为了避宋太宗赵光义的讳而改为宜兴。再比如有的地方以皇帝的年号命名，比如陕西淳化县就是以宋太宗的年号淳化为名。如果不了解这些历史，就不能准确理解地名的含义。地名词语还能告诉我们许多民族史的知识，例如江浙一带的地名中有以“于”、“无”、“余”、“句”、“姑”打头的地名，如于潜、无锡、余杭、句容、姑苏等，这些字都是古越语的发语词，没有实在意义，这类地名证明了江浙一带为古越族的活动范围。至于地名源流的研究，那就更需要用历史学的方法进行研究。可见，地名学研究也离不开历史学。

总之，地名学是和语言学、地理学、历史学、民族学、测绘学等多种学科密切联系的一门综合性学科。

3. 汉语地名研究

我国的汉语地名研究源远流长。《说文解字》、《释名》和《尔雅》当中都有丰富的地名资料。比如《尔雅》中对于不同的自然实体和地域就做了初步

的划分。《尔雅·释地》:“邑外谓之郊,郊外谓之牧,牧外谓之野,野外谓之林,林外谓之埛。”这段话说明了以城市为中心的四周地域分布的差异。《尔雅·释水》:“水注川曰溪,注溪曰谷,注谷曰沟,注沟曰浍,注浍曰渎”,“江河淮济为四渎,四渎者,发源注海者也。”这段话说明了对河流通名的分类。

班固的《汉书·地理志》始创正史中的地理志,开了地理总志的先河。书中收录地名达4 500条以上。《汉书·地理志》首先叙述全国的山水物产,然后在郡名下注明所属州,在郡名后列有居民人口。许多地名还讲述了设置的来历,名称的更改,境内的名山水道、矿产资源等。北魏郦道元的《水经注》收集了2万条左右的地名,详细阐述的有2 000多条。以水为纲,记载了各地的名山,还概括了地名命名规律。唐代李吉甫的《元和郡县图志》是一部全国性的地图和地理总志,极富地名研究价值。

但是,我国古代的地名研究局限于文字训释和材料的罗列,地名学还没有形成独立的学科。现代地名学的形成,始于20世纪30年代。1933年出版了《中华民国新地图》,随后出版了臧励和等编纂的《中国古今地名大辞典》。研究历史地理的禹贡学会也在此时成立,《禹贡》半月刊开始发行。

解放后,我国地名学取得了很大的成就。1959年曾世英在测绘研究所建立了地名研究室,为清理外国侵略者强加给我国的地名、维护国家主权作了大量的工作,为少数民族地名译写和中国地名的国际标准化作出了重大贡献。1977年,国务院成立了中国地名委员会。1979年召开全国首次地名工作会议以后,全国各地展开了地名普查。1981年开始组织专家和地名工作者编写《中华人民共和国地名词典》。1988年成立了中国地名学研究会。经过20多年的发展,汉语地名学逐渐成熟,出版了大量的、高质量的地名学论著,无论是在理论研究还是实际应用,都取得了可喜的成绩,为我国的现代化建设作出了应有的贡献。

## 二、人名和人名学

### 1. 人名

(1) 人名的属性

人名有广义和狭义之分,广义的人名包括姓和名,狭义的人名不包括姓,只指名。我们所讲的人名是广义的人名。

人名是指称个人的专有名词,是代表一个人的语言符号,所以人名首先

是一种语言现象，具有语言属性。任何人名都是以语言文字为物质材料而存在的，每个人名都有自己的语音形式、书写形式（有文字的话）、内部结构和指称对象等基本的语言特征。比如“张飞”这个人名，语音形式是“Zhāng Fēi”，拼音书写时姓和名的第一个字母要大写，用汉字书写为“张”和“飞”两个汉字；“张”是姓，“飞”是名，用来指称三国时代的一个将军。因此，人名和地名又构成语言学的一个分支——专名学。

人名因适应社会生活而存在，只要进入生活和社会，一个人便会产生名字，社会属性是人名的基本属性之一。人名的社会性主要表现为具有约定俗成的指别性、专一性和永久性。人名最基本的功能就在于称呼、指别作为个体存在的人。比如在一个小群体当中，如果有重名，我们就不得不使用一些附加的特征来区分不同的人，比如“大张明”、“小张明”等等，指别性是人名的基本属性；一篇文章，署上作者的姓名，表明这是作者的研究成果；一项专利，注上发明人的姓名，表示该专利为发明人所有。在现代社会，公民还享有姓名权，姓名受到法律保护，《民法通则》规定：“公民享有姓名权，有权决定、使用和依照规定改变自己的姓名，禁止他人干涉、盗用、假冒。”这是人名的专一性；一个人有了姓名之后，姓名要陪伴他走过一生，一个人的名誉、权力、人格等等都凝聚在人名上。有的人流芳百世，比如岳飞、文天祥、李白、杜甫、张衡、祖冲之；有的人遗臭万年，南宋奸臣秦桧残害忠良，后人把秦桧夫妇的跪像立在西湖岳飞墓前，墓前有一副广为流传的对联：“人从宋后羞称桧，我到坟前愧姓秦。”所谓“树的影，人的名”、“行不改名，坐不改姓”都是指人名的永久性。

中国人和外国人的姓名不同，汉族和少数民族的姓名也不一样，人名还是一个民族文化传统的反映，文化属性是人名的重要属性。中国古代是一个宗法专制社会，中国人重视家族，家族和家庭比个人占有更为重要的位置，因而中国人特别看重表示血缘、祖根、家族的姓氏，不是迫不得已，一般不会更改自己的姓氏。在命名上，也十分重视辈分，不同辈分的人使用不同的辈字，体现出了前后相继的关系。以孔子家族为例，从元代孔子第 54 代孙孔思晦开始，辈字依次是思、克、希、言、公、彦、承、弘、闻、贞、尚、衍、兴、毓、传、继、广、昭、宪、庆、繁、祥、令、德、维、垂、佑、钦、绍、念、显、扬。全国政协委员中有 3 位是孔子的后裔，他们是孔祥祯、孔令朋和孔德懋，根据辈字，很容易知道他们分别是孔子的第 75、76、77 代孙。有的辈分用字简直就是一首诗，比如张之洞规定 20 个字作为子孙的辈行：

仁厚遵家法 忠良报国恩
通经为世用 明道守儒珍

宗族思想在名字中也有表现，比如怀祖、念慈、孝先、家和、宗仁、乐群、广嗣、延嗣、裕孙、世隆等等。因为尊祖敬宗，中国还存在特有的避讳现象。对帝王的名字要避讳，比如“正月”的“正”读平声，不读去声，这是因为秦始皇名嬴政，和“政”同音的“正”就改读阴平。汉明帝名庄，为了避讳，把庄姓改为严姓，《庄子》也就成了《严子》。对长亲的名字也要避讳。司马迁的父亲叫司马谈，《史记》中就把“赵谈”改为“赵同”，“张孟谈”改为了“张孟同”。大诗人杜甫的母亲名字叫做“海棠”，所以杜甫的近 3 000 多首诗歌中没有一处提到海棠，正如苏轼所说：“恰似西川杜工部，海棠虽好不留诗。”

从人名自身的结构来看，人名又具有系统性。原始人类是以氏族为单位从事生产活动的，上古的名号，大多是氏族公用名，还没有产生代表个体的私名。有巢、燧人、伏羲、神农等既是氏族首领的名号，也是这个氏族所有成员的公名。代表个人的私名的出现，应该在私有制出现之后。私有制出现以后，个人的财产、权益需要维护，自我从公有的氏族中脱颖而出，于是，私名的建立成为需要。在私名的建立过程中，大多数人名系统经历了私名加公名的时期，公名逐渐演变为姓氏。随着社会的发展，私名在形式上逐渐复杂化，出现了联名制。比如阿拉伯民族的姓名由个人名、父亲名、祖父名、别名和称号五部分组成，俄罗斯人名由个人名、父亲名和姓三部分组成，此外还有小名、爱称、简称、尊称等变体。中国的人名结构形式很简单，由姓加名组成，但是，中国人名的变体很多，名、字、号、谥、地望名、官爵名、笔名、绰号等构成了汉族人名独特而复杂的系统。

(2) 中国人名的基本特点

中国人名形式繁多，内容丰富，概而言之，中国人名的基本特点表现在以下几个方面：

第一，汉语是中国人名的载体，汉语的特征决定了中国人名的形式特点。汉语是有声调的语言，元音占优势，音节结构简单，讲究韵律的和谐。这些特点使得汉语人名具有很强的音乐性。早在先秦时代就出现了双声、叠韵的人名，如胡亥、扶苏，后代就更多了，比如扬雄、曹操、杜甫、杜牧等。现代更有许多叠音的人名，比如方芳、魏巍、杨阳。汉语人名的长度也受到汉语语音特点的制约。汉语是讲究节律的语言，双音步是汉语的基本的韵

律结构，因而汉语人名以两个和3个字为最多，最多4个，超过4个就拗口累赘，称呼起来很不方便。即使是3个字的人名，人们也喜欢用两个音节来称呼，如果是单姓，就称双名；如果是复姓，就乐于用姓来称呼。

第二，多名制度。在当代社会，中国姓名制度大大简化了，一个人一般只有一个名字。但是，在古代中国，实行的是多名制度。一个人不同的人生阶段有不同的人名，所谓幼有名，长有字，死有谥。除了名和字外，古人还有别号，别号为自己所取，体现了不同的追求和个性。进入社会后，由于社会地位的原因，又有了地望名、官爵名；由于所从事的职业，艺人又有了艺名，僧道也有法名和道号。此外，封建帝王还有庙号、年号、尊号等称呼。所以，古代中国的人名制度相当复杂。

第三，从文化内涵上看，中国人名具有浓厚的礼教色彩。君君臣臣，父父子子，封建礼教渗透到人名系统的各个角落，重姓轻名，讲究辈分，为尊亲避讳等等人名现象都体现了这一点。可以说，人名系统是中国传统文化的一面镜子。

**2. 汉语人名学**

人名学是研究人名系统的起源、发展、结构及其内涵的一门学科。人名是人类社会发展的产物，它与人类的语言、历史、文化、社会都有密不可分的关系，人名系统的背后蕴藏着一个民族的信仰、习俗、价值观、审美观、社会心理、时代风貌等丰富的内涵。研究人名，首先要搞清名字的含义及变体等问题，这就必须借助语言学的知识；探讨人名系统的发展变化，又离不开历史知识的支持；研究人名的结构，必须对一个社会的发展、组织状态有所了解，这又离不开人类学；各个民族的人名系统有很大的差异，如果没有民族学的知识，要深入理解一个民族的人名系统简直寸步难行。所以人名学从产生的那天起，就是一门综合性的边缘学科。

中国姓名的研究可以上溯到战国时代的《世本》。《世本》记载了从黄帝到春秋的帝王卿大夫的世系、传谱等，不过原书已经散佚。《尚书》、《春秋》、《左传》等先秦典籍都有关于姓氏的记载，其中《左传》提出了姓氏的6种来源，还记载了大夫申缙的一套命名理论。现存较早的姓氏著作是西汉史游所撰的《急就篇》，该书共罗列姓氏130个。东汉王符的《潜夫论》有"志氏姓"一章，从历史的源流讨论了姓氏的来由。应劭的《风俗通义》更说明了500多种姓氏的来源。班固的《白虎通·姓名》还谈到了名和字的关系。许慎从文字的角度对姓的解释很有价值，他还初步探讨了姓氏的发展。魏晋

以来，门阀制度盛行，与此相适应，以谱牒学为主的姓氏学也发达起来。代表作有东晋贾弼的《姓氏簿状》，该书有700多卷，划分18州、116郡的世家大族的姓氏等第。唐代官修了《大唐氏族志》100卷和《元和姓纂》10卷，为后世提供了丰富的姓氏材料。宋代是姓氏研究的重要时期，代表论著有郑樵的《通志·氏族略》、邓名世父子的《古今姓氏辩证》和王应麟的《姓氏急就篇》。郑樵对我国的姓氏起源、演变及其研究的发展作了精辟的阐述，考察了27种姓氏的来源，并涉及了2 255个姓的源流，他对于姓氏的看法也有很大的影响。明代凌迪知所撰的《万姓统谱》104卷，较为全面；还有杨慎《希姓录》和《希姓存参》两部专述希姓的著作。清代姓名著作以考据为特色，顾炎武的《日知录》、钱大昕的《十驾斋养新录》和俞樾的《春在堂随笔》堪为代表。王引之从训诂学的角度作《春秋名字解诂》，精辟地说明了名和字的义类五体：同训、对文、连类、指实、辨物，开创了名字训诂学。综观19世纪以前的人名研究，有两个显著的特点，从研究的对象上来看，偏重于姓氏的渊源考释；从方法上说，重视资料的收集整理，单一地采取语言学的方法，缺少理论和研究方法上的突破。

进入20世纪，姓名研究出现了新气象。王国维的《殷商制度论》，把考古学、历史学和人名学结合了起来，否定了中国姓氏制度始于远古时代的结论，认为周代才建立了"男子称氏，女子称姓"的制度。30年代初潘光旦的《中国家谱略史》，较为全面地说明了历代的谱学作品及它们的编纂、修订的情况。30年代中期，袁业裕编著《中国古代氏姓制度研究》，是一部系统、全面论述姓氏制度的专著。尤其值得注意的是，从30年代开始，一批民族学家、语言学家、历史学家对我国的少数民族的联名制展开调查，开了综合研究人名的先河。比如罗常培的《论藏缅族的父子连名制》、杨希枚的《联名与姓氏制度的研究》等论著。80年代后，人名研究出现了新高潮，出版了一系列人名学著作，比如徐俊元、张占军、石玉新合著的《贵姓何来》、王泉根的《华夏姓名面面观》、萧遥天的《中国人名的研究》、袁玉骝的《中国姓名学》、王建华《文化的镜象——人名》、张联芳主编的《外国人的姓名》和《中国人的姓名》、纳日碧力戈的《姓名论》，中国的人名学走向了成熟。

人类已经进入了21世纪，随着信息化的来临，时代向人名学提出了新的任务，人名的规范化、标准化成为迫切需要研究的课题。网络时代的来临，也给人们对人名的理解、制定和应用提出了新的挑战，相信人名学一定会为社会的发展作出更大的贡献。

# 第二节 汉语地名学

## 一、地名词语的语言学特征

地名是人们共同约定的一种语音符号,属于语言范畴。它必须借助于语言和文字来表示。因而,地名词语包括三要素,即地名的含义、记载地名词语的字形及地名词语的读音和意义。但是,在形、音、义上,地名又有不同于语言中普通词汇的特征。那么,地名在语音、字形和意义上有什么特点呢?

### 1. 地名词语的含义

(1) 指称意义

地名都是因地命名或为地命名的专用名称,它和一般地理名词不同。一般的地理名词虽然也是以客观存在的地理实体为命名对象,但是更着眼于"类"的划分,具有很大程度的概括性,例如"山脉"、"河流"等等,都是所有具体山、河等的泛称,而地名是一个一个地给地理实体命名,所以,地名词语最基本的含义首先是它指代的地域位置、范围和类别。

一个汉语地名词语一般由通名和专名两部分组成,通名多为一般的地理名词,而专名范围较广,任意性较大。例如"黄河"这个地名,我们从通名"河"识别了它代表的地理实体是一条河流,而不是山脉、平原、村落;从专名"黄"同别的河流如长江、淮河、珠江等区分开来,从而确定了"黄河"这个地理个体的性质、种类,同时也确立了这个地理实体的方位和范围。

当然,地名所代表的方位有自然的,也有人文的。例如行政区划,虽然行政区划范围的确定不一定具有地形、水文、植被等自然地理的特征,但是它也代表了一定的地域范围和特征,只不过属于人文地理而已,例如一说"陕西省",那么陕西省所处的特定的方位、范围就确定了。

总之,地名是人们约定俗成的用来指称具有一定方位、特定范围地域的名称,它的作用就是把千千万万个地理区域区分开来,因而地名的指称意义是地名词语最基本最核心的含义。

(2) 命名意义

地名词语的命名意义体现了地名命名的理据和来源。很多地域的命名是以地理特征为理据命名的。比如:山西(太行山之西)、河北(黄河之北)、

湖南(鄱阳湖之南)、洛阳(洛河之南)、华阴(华山之北)、渭南(渭水之南)等是以地理方位命名的;黑龙江、浙江、青海、洪泽、湘潭、牡丹江等是以水命名的;平顶山、五台山、马鞍山等是以山命名的;五指山、葫芦岛等是以地形特征命名的;榆林、鸟岛、蛇岛、攀枝花等是以生物命名的;铁岭、盐池、金沙江等是以岩矿资源命名的。

还有一些地名或者是以居民姓氏命名的,比如张家庄、王家坟、辛家庙;或者是以历史人物命名的,比如中山(市)、左权(县)、志丹县;或者以人们的愿望而得名的,比如西安、丰宁、承德、怀柔、长治等等。

还有的地名涉及社会历史,例如"宝鸡"是陕西省的一座历史名城。宝鸡原名陈仓,因陈仓山而得名。唐代安史之乱爆发后,唐玄宗奔蜀,肃宗于战乱中在甘肃灵武即位,进驻西安后改陈仓为宝鸡,取的是"宝鸡殷鸣"的吉祥之意。

(3) 附加意义

地名的基本意义是指称意义,但是有些地名由于环境或人为特点而获得的知名度大大超过了这些地名的所指称的地域位置。例如人们都知道黄果树瀑布,但是未必知道黄果树瀑布的具体位置,人们都知道翠亨村是孙中山先生的家乡,但未必会知道翠亨村具体的地理方位,说起珠穆朗玛峰,我们都知道它是世界最高峰,反而不大在意它的具体方位,所有这些,都是地名的附加意义。显然,每一个地名都具有指称意义,但并不是每个地名词语都具有附加意义。

**2. 汉语地名词语的结构**

汉语地名词语的结构可以从不同的角度来看。从结构成分的指称特点来看,可以分为专名和通名;从造成词的性质来看,可以分为原生地名和派生地名;从语法结构来看,可以分为单纯词地名和复合词地名。下面分别讨论。

(1) 通名和专名

标准化地名由专名和通名两部分组成。例如"山西省"、"北京市"是标准化地名,"山西"、"北京"是专名,"省"、"市"是通名。专名是识别个体名词的很重要的标记,通名是识别个体地名所属类别的标志。在日常应用中,通名经常省略,比如"山西省"、"北京市"经常说成"山西"、"北京",但是在汉语中,单音节的专名和通名不能单说,两者必须结合在一起才能称说,比如"通县"、"李庄"、"长江"、"黄河",专名"通"和"李","长"和"黄"是单音节词,通

名“县”和“庄”、“长”和“黄”就不能省略，必须说全称。此外，指称自然地理实体的地域名称，一般也不能省略其通名，如黑龙江、雷州半岛中的“江”、“半岛”就不能省略。

从历史上看，早期的地名没有通名。比如“河”和“江”在古代是黄河和长江的专名，后来变成了通名。北方的河流多用“河”，比如“海河”、“淮河”等，南方的河流多用“江”，比如“珠江”、“黄浦江”。这是专名转化为通名的性质。

通名转化为专名的情况也较为常见，比如“秦皇岛”，“秦皇”是专名，“岛”是通名，“秦皇岛市”，“秦皇岛”成了专名，“市”为通名。再如“河市”，“河”是专名，后边的“市”为通名。

(2) 原生地名和派生地名

地名往往不是孤立的，常常会有地名结群的现象。例如北京城区，以广安门为原生地名，派生出广安门内大街、广安门北滨河路、广安门外大街、广安门南街、广安门北路、广安门南路等。

原生地名派生出派生地名，主要有以下几种类型。一种是同地派生，有的是合二为一。例如安徽至清康熙六年设安徽历政使司，以巡抚驻地安庆府(今安庆市)和徽州府(今歙县)两地首字为省名。再如江苏，是取江宁(今南京市)和苏州的首字为省名，福建取福州和建宁两地首字为省名。也有的是一分为二，比如西安的长乐路分为长乐东路、长乐西路和长乐中路。辛庄分为前辛庄和后辛庄。

另一种是异地移用而派生的地名。这种情况以上海最为典型。上海城市道路很多，移用了许多省市名化为道路名称，比如西藏路、北京路、南京路、延安路、陕西路、桂林路，等等。

(3)汉语地名词语的语法结构

1) 单纯词地名。单纯词地名多是历史地名。古代汉语是单音节为主的语言，因而最早产生的汉语地名多为单音节单纯词构成的，它们都是专名。例如秦、齐、燕、韩、赵、魏、楚、鲁、宋，这些历史地名，有的成了某些省的简称，如陕西省称秦，山西省称晋，山东省称鲁，有的加通名沿用至今，如山东的莒县、曹县、滕县，河南的杞县、邓县、密县、息县、温县，山西的祁县、蒲县等。

还有的地名是双音节的单纯词，有的是双声，如盱眙。有的是叠韵，如零陵、邯郸、昆仑。还有的是既非双声，也非叠韵，如敦煌、琅琊。

2）复合词地名。复合词地名是汉语地名的主体，它的构造方式和现代汉语构词模式相同，主要合成方式有偏正、附加和联合3种。

① 偏正式。偏正式是复合地名的最主要的方式。它的形式是专名在前，通名在后，又可以细分为以下几种情况。

一种是名＋名。如：徐州、杨店、石塘、柳林、花巷、渭河、华山、潼关、珠江。

一种是形＋名。如：赤壁、赤水、长江、白水、黄山、白塔、香港、清涧。

一种是方位词＋名。如：南京、北京、西安、上海、下关、中山、东湖。

一种是数量词＋名。如：五台山、二十里铺、三岔口、万县、万泉河。

② 附加式。构成方式是在通名后加上方位词，例如：山东、山西、河南、河北、江西、关中、海南。

③ 联合式。主要由几个较小的地名缩略而成，例如：江浙（江苏、浙江）、云贵川（云南、贵州、四川）、陕甘宁（陕西、甘肃、宁夏）、杭嘉湖（杭州、嘉兴、湖州）。

3）词组地名。随着社会生活的发展，聚落的复杂化，人们需要更为准确、标准的地名，因而出现了大量的层层套叠限制的词组地名词语。例如"三江源自然保护区"这个地名词语，"三江源"是专名，"自然保护区"是通名。从语法构造上来说，是一个偏正结构词组。"三江源"限制"自然保护区"，指明是"三江源"而不是其他自然保护区。"三江源"本身又是一个偏正结构，"三江"是缩略的专名，"源"又是一个通名，进一步限制了所指的地域范围。这个名称不仅指出了地域范围，还指明了这一范围的性质（自然保护区），给人们提供了充足的信息。

在现代社会，标准地名大都是层层限定的多级地名，因为一个小的地域总是归一个大的地域，只有层层限定，才能够准确地确定一个地方的准确方位。以通信为例，"北京"是中国的首都，知名度很高，但是从国外寄往北京的信件，也要加上"中国"作为限定词。至于小地名，如果没有层层限定，那是没有办法传递的。因而，在现代社会，多层的词组地名是非常重要的。

**3. 地名词语的语音用字特点**

(1) 地名中保存着较多的古音古字

语音是发展变化的，一般口语里的字音随着语音发展变化而变化，但是地名里的常用字是比较固定的，它的读音世代相袭，保留着较多的古代读音。比如"番"在口语中读 fān，比如"番茄"、"番薯"。但在广东省地名"番禺"当中读为 pān。"费"在口语中读为 fèi，但在山东省地名"费县"中读为

bì。它们在地名都不读唇齿音 f(古代称轻唇音),而读双唇音(古代称重唇音)b\p。这是上古音的保留,音韵学上叫作"古无轻唇音"。再比如"铅"在一般口语中读作 qiān,但在江西的"铅山县"中读作"yán",这也是古音的保留。"铅"在古音是以母字,是零声母,在地名中保留着古音。

同样,在用字方面,地名词语也具有保守性,有的地名一直沿用着古代的地名用字,比如安徽的亳(bó)县、黟(yī)县、歙(shè)县、砀(dàng)县;再如陕西的醴(lǐ)县、郃(hé)阳县、鄠(hù)县、邠(bīn)县、鄠(hù)县、沔(miǎn)县、汧(qiān)县。这些生僻字难认难写,为了方便使用,在《简化字总表》中对这类字进行了简化,一般采用同音代替的方法,比如醴泉改为礼泉,汧阳改为千阳,邠县改为彬县,鄠县改为户县。

(2) 地名中多方言用字和方言读音

我国的地形复杂多样,许多地名是因为地貌而得名,有些地貌是某一地区特有的,因而,不同的方言区就有了自己特有的地貌通名,也就产生了独特的用字,比如黄土高原为陕西、山西特有的地貌,因而秦晋两省有许多以"塬"、"梁"、"峁"、"砭"等地命名的地名。比如:

"塬"是指黄土高原上四周为流水切割,顶面广阔的地貌,"塬"是良好的耕作区,陕北有洛川塬,陕西华阴有孟塬。

"峁"是黄土丘陵,顶部浑圆,斜坡较陡,以"峁"得名的有陕西米脂的桃花峁,陕西神木的沙峁。

"梁"是黄土高原上条状延伸的山岭岗,因此得名,如陕西定边堆子梁,山西沂州的朱家梁。

"砭"是指黄土高原上傍山临沟的狭长通道,延安的"青化砭",著名的青化砭战役就发生在这里。

再如长江三角洲河网重布,就产生了溪、泾、浦、汇等与江河有关的吴方言用字,比如上海有"陆家溪"、"青浦县"、"徐家汇"、"同浦"。

一个地方的地名,大多数最初是由"主人"来给命名的,不难想像,在命名之初,大多数地名就是用方言来命名的,本地人可以用方言来称说,需要用标准音称说时,就像方言中的其他词语一样,方言词语的读音可以折合为相应的标准读音。但是,有的方言词命名的地名比较粗俗,于是有的文化人就用读音相近的文雅汉字来书写,这样就形成了"雅俗异名"的现象,比如北京有很多含有尾巴、肉、毛、皮、猪等鄙俗字眼的胡同,就改成了雅名,例如:

羊尾胡同——扬威胡同　　驴肉胡同——礼路胡同

| | |
|---|---|
| 猪市胡同——珠市胡同 | 灌肠胡同——官场胡同 |
| 臭皮胡同——受壁胡同 | 母猪胡同——梅竹胡同 |
| 干鱼胡同——甘雨胡同 | 鸡爪胡同——吉兆胡同 |

(3) 地名中的异读和异写

地名词语当中存在着不少异读现象。有些异读是属于多音多义字，比如“华家屯”以姓定名，读音为 huà；“中华路”以我国国名定名，读音为 huá；类似的，四川有“乐(lè)山”，浙江有“乐(yuè)清”；云南有“大(dà)关县”，河北有“大(dài)城县”。还有的是因为保留古音而造成的异读，如“蚌”在“河蚌”中读 bàng，在安徽的“蚌埠”中读 bèng；或者是因为不同地方的习惯读音，比如“侯”在河南侯马中读为 hóu，在福建闽侯中读为 hòu。

还有的是不同来源的通名使用了同一个汉字，造成了异读。例如，“堡”读 bǔ 时是指有墙壁围护的村镇，如陕西东北部的吴堡县，山西代县南的阳明堡；在南方，这个意义读 bǎo，如南京有“中堡”。“堡”还读为 pù，同“铺”，是指旧时的驿站，比如十里堡、三十里堡。也写作“铺”，如十里铺、三十里铺。

地名中，尤其是小地名中还存在着很多异写。比如圪垯是一种圆形的小土丘，因为各地的读音不同，因而写成了不同的字，比如灵宝写作疙瘩，封丘写作葛塔，还有写作圪旦、圪堆、圪垛等。

## 二、地名词语系统

世间地名千千万万，五花八门，本身存在着种种差异。昆仑山指称山脉，黄土高原指称高原，四川盆地指称盆地，长江、黄河、太湖、太平洋指称水体，台湾岛、橘子洲指称岛屿，北京市、江苏省指称行政区域，天安门广场、颐和园指称人文建筑地名。但是，地名之间也存在着众多的共同之处，比如山脉、高原、盆地、岛屿，都是地形面貌的名称，长江、黄河都是水系名称，可以把它们进一步概括为自然地名；省、市是表示城乡聚落的，和建筑地名一样，都是人文地名。可见，地名既有差异性，又有共性，地名总是处在大大小小的系统当中，这是我们描写地名词语系统的基础。

地名词语系统可以从不同的角度划分，鉴于地名词语本身由通名和专名构成，我们准备从通名和专名的角度来描写地名词语系统。

### 1. 从通名的角度划分地名词语系统

从通名的角度来刻画地名词语系统，依据的是通名所反映的个体地域属性。一般来说，通名不同，它们所指称的个体地域的属性就不同，比如四

川盆地、黄土高原、东北平原中的通名分别是“盆地”、“高原”和“平原”，这3个通名指称了不同的地形；通名相同，它们所指称的个体地域的属性也基本相同，比如京广线、京沪线、京哈线，都是铁路线，具有相同的通名。一般把通名系统归纳如下。

(1) 自然地域名称

1) 陆地地域名称。根据所指地域的地理特征，又细分为以下几种。

大陆地名：欧洲、亚洲、非洲、南美洲。

平原地名：长江中下游平原、华北平原、珠江三角洲、长江三角洲。

高原地名：黄土高原、云贵高原、内蒙古高原、青藏高原。

盆地地名：四川盆地、柴达木盆地。

沙漠地名：塔克拉玛干沙漠、毛乌素沙漠。

丘陵地名：山东丘陵、江南丘陵。

草原地名：呼伦贝尔草原、松潘草地。

山脉地名：喜马拉雅山脉、大兴安岭、秦岭、五指山、太行山、华山。

半岛地名：山东半岛、雷州半岛。

2) 水域名称。

海洋地名：太平洋、北冰洋、印度洋、黄海、东海、南海。

江河地名：长江、珠江、黄河、海河、汉水。

湖泊地名：太湖、洞庭湖、白洋淀、荷花淀、滇池。

海峡地名：台湾海峡、直布罗陀海峡。

海岛地名：台湾岛、海南岛、钓鱼岛、澎湖列岛。

(2) 人文地域名称

1) 行政区划名称。

省级行政地名：江苏省、北京市(中央直辖市)、宁夏回族自治区。

地级行政地名：南京市(省辖市)、关中地区、大理白族自治州、哲里木盟。

县级行政地名：玄武区(市辖区)、义乌市(县级市)、神木县、丰宁满族自治县、杭锦旗。

乡镇行政地名：后子头乡、大柳塔镇。

2) 聚落地域名称。

城镇地名：西安城、北京城、凤栖镇、吴起镇。

街巷地名：王府井大街、南京路、东交民巷、高家弄。

小区地名：杜鹃园、龙江小区、清江花苑。

村落地名：平阳村、张家庄、十里铺、牛堡屯、高家井、胡家庙、太和寨。

3）人工地名名称。

交通设施地名：京沪线、北京西站、宁镇公路、210国道、三环线、(上海)明珠线、虹桥机场、董家渡、大连港、南浦大桥、南京长江大桥。

水利设施地名：都江堰、三峡大坝、红旗渠、荆江大堤、苏北灌溉总渠、京杭大运河、丰满水电站。

文化设施地名：故宫博物馆、山顶洞人遗址、颐和园、大雁塔、敦煌石窟、杨家岭毛泽东故居、遵义会议旧址、兴庆公园、红山动物园、上海图书馆、清华大学、工人体育馆。

企业厂矿地名：大庆油田、宝山钢铁厂、大兴安岭农场、舟山渔场。

***2. 从专名的角度划分的地名词语系统***

地名的专名反映了地名命名的依据和方式，从专名的角度分类，实际上就是对汉语地名的语源进行分类，一般归纳为如下类别：

(1) 因水得名

1）以河流而得名。

河北省、河南省、黑龙江省、浙江省、四川省、松江(上海市)、漠河、涞水县、赤水市、洛川县。

2）以湖泊而得名。

湖北省、湖南省、青海省。

3）以泉、池、井而得名。

甘泉(陕西)、酒泉(甘肃)、泉州(福建)、甘井(陕西)、四眼井(内蒙古)、渑池(河南)。

4）以海而得名。

威海市、上海市、海口市。

(2) 以地得名

1）以山得名。

鞍山市、平顶山市、崂山县、岐山县、嵩县(因嵩山得名)、赤峰市、黄冈县、曲阜市、孟良崮。

2）以地理位置而得名。

山东、山西、南京、北京、河南、河北、咸阳、江阴、华阴、汉口、临汾(地临汾水)、临河(地临黄河)。

3）以地物形态特征而得名。

日月潭、葫芦岛、长白山、五指山、月牙泉、火焰山。

(3) 以物产而得名

榆林、枣园、盐山、盐城、海盐、铁岭市、铜川市、金沙江、石棉县、油泉子。

(4) 以人得名

中山市、左权县、志丹县、张自忠路、张家庄、李村。

(5) 以史得名

1）沿用古国名。

山东的曹县、滕县、费县，河南的杞县、邓县、巩县、温县、滑县、虞城县，陕西的户县、彬县，山西的祁县、蒲县、绛县等，都是沿用先秦时期的地名。

2）以封建帝王的年号得名。

景德镇原名昌南镇，宋真宗景德年间改名。绍兴原名山阴，宋高宗绍兴年间改名。类似的还有陕西的淳化县（宋太宗淳化四年），福建的永泰县（唐代宗永泰二年）、政和县（宋徽宗政和五年），江西的兴国县（宋太宗太平兴国七年），河南的登封县（唐武则天万岁登封元年）。

3）以历史事迹而得名。

山东的烟台市，在明朝是海防重地，设烽堠，防倭寇，所以叫作烟台。辽宁的旅顺这一名称始于明朝洪武年间，明朝经营辽东，士兵由山东蓬莱渡海，取旅途顺利的意思。山西的闻喜县和汉武帝有关，元鼎六年汉武帝至此，闻南越攻破，所以设置了闻喜县。

4）因为避讳而改名。

浙江的龙泉县，原名龙渊县，唐代避唐高祖李渊的讳，改称龙泉县。江苏的宜兴、陕西的宜川，原名为义兴、义川，因为宋代避宋太宗赵光义的讳，而改为宜兴和宜川。

(6) 以部族或民族得名

四川的巴县、巴中，湖北的巴东，都是先秦巴族聚居区，因此得名。春秋时越族建立吴（勾吴）、越（於越）产生了以勾、於冠首的地名，如句容、姑苏（句、姑为勾、於的转写）。

(7) 因复合而形成的地名

安徽是安庆和徽州首字的合称，淄博是淄川和博山的合称，武汉是武昌、汉口、汉阳三地的合称。

(8) 以交通而得名

天津原名直沽寨，后改名为海津镇，明成祖率兵从海津镇渡卫河南下进京师，取得帝位后便把海津镇改为天津，意为“天子的津渡”。河南有孟津，孟是第一的意思，津是渡口的意思，相传是夏禹治水时从三门峡到此建立的第一个渡口，所以叫孟津。类似的还有四川的渡口市、贵州的乌江渡、赤水的太平渡、湖北的仙人渡。再如“铺”旧为驿站名，很多地名以“铺”为通名，例如陕西有四十里铺、江西有大路铺。再如北京的“通县”取义“漕运通济”，甘肃的通渭县位居渭河支流，由此可以通向渭河平原，所以叫“通渭”，与此相类似，内蒙有通辽市，甘肃有通北口，都因交通得名。

(9) 以经济活动而得名

河南的牛市屯、四川的桔子铺、广西的黄牛市，以交易物命名，说明这些地方以前是农牧产品贸易的集市。城市里也不乏其名，比如炭市街、米市大街、骡马市大街、菜市口等，更多的是以工业命名的地名，诸如石油大街、矿山路、纺织路等等。

(10) 取吉祥字眼命名

用吉祥的字词取名，表达了人们美好的愿望。比如陕西有西安、长安，甘肃有安西，山西有长治，浙江有吉安，广西有南宁，河北有永定河。

(11) 借用地名

借用省市名称给城市的道路命名是借用地名的最常见的方式。比如，上海市移用了我国许多省市名称作为市区道路的名称，如南京路、陕西路、延安路、桂林路、西藏路，等等。与此类似，南京也有上海路、北京路、山西路，等等。

地名的借用，并不是杂乱无章的。以上海为例，上海把最繁华的商业街叫做南京路，是因为解放前南京是中国的首都；为方便起见，上海市区外围以其他省市命名的道路，大体上和这些地名在我国版图上的位置对应，东北部多以东北三省的地名作为道路名，比如延吉路、抚顺路等，西北部多以西北的地名作为道路名，比如延安路、陕西路、新疆路等。

## 三、地名规范化

地名随着人类社会的产生而产生，随着人类社会的发展而发展。人们天天说地名，用地名，衣食住行都离不开地名，和人交流离不开地名，科学研究也离不开地名。地名为人们的各种活动创造了条件，是人类活动的基础条件之一。人类的生活需要科学的地名学，地名学也只有在现实生活的应

用中才会取得发展。地名学的应用很广，但是，地名标准化是地名学应用的首要目标。

地名起源很早，远古时期就有了。那时的地名，数量比较少，也比较简单。随着社会的发展，人类活动范围的扩展，地名就越来越多了，也越来越复杂了。我国历史悠久，幅员辽阔，地名更是繁杂，存在着很多问题，最突出的有以下几方面。

一是一地多名。我国历史悠久，地名时有变更。虽然有了新名，但是旧名还在习用。比如西安在历史上长期称为长安，在清朝末年，西安城中还并存着西安和咸宁两县，后来咸宁废县，长安外迁，西安才成为定称。这是古今地名的差别。北京的许多胡同都有俗名、雅名的差别，俗名和雅名大多谐音，但是雅名用字文雅，比如“驴市街”是俗名，后来写成“礼士胡同”，“闷葫芦罐”是俗名，后来写作“蒙福禄馆”。这是俗名、雅名的差别。

二是一名数地。比如吉林省和吉林市的专名相同，吉林医学院在吉林市，而吉林工学院在长春市，信件邮寄很容易出错。如果从口语交际的角度看，同音地名也会引起混乱。比如澧县—礼县—理县—蠡县从字形上看是4个不同的地名，但是读音都是Lǐxiàn，没有区别。

三是地名生僻用字多，异读异写多，不方便交流使用。这在上文已经讲过，兹不赘述。

这些地名使用当中的混乱现象，在相对闭塞的古代社会，或许可以容忍，但在交往越来越频繁的现代社会，地名使用不科学，就会造成很大的不便，甚至引发事故。国内如此，国家与国家之间的交流也是一样。如果没有规范标准的地名，不难想像，各种国际交流就会遇到极大的障碍。因此，做到一名一地，一地一名，字有定写，音有定读，实现地名标准化，是地名工作的首要目标。具体来说，要做到：

第一，分化重名。经过20世纪80年代的地名标准化处理，重名问题得到了较为全面的清理。但严格地说，地名中还有许多同音不同字的地名，它们实际上也是重名现象，因为写成拼音后就没有了区别，比如：

| | |
|---|---|
| Chángníng | 长宁，常宁 |
| Bǎoshān | 宝山，保山 |
| Yìyáng | 益阳，弋阳 |

如果拼写时没有声调，问题就更多了，例如：

Ningjin　　宁津，宁晋（河北）
Pinglu　　平鲁，平陆（山西）

如何在拼写上区分它们，还需要进一步研究。

第二，规范地名用字。地名所具有的历史继承性和区域性，使得地名中存在着许多生僻字。这些字难认难读，好多字在词典里也没有收录，它们的存在，给应用计算机录入和检索地名，都造成了很大的不便。建国以来，根据便于认读的读音相近的原则，对一批地名生僻字进行了更改，比如陕西的鄠县改作户县，新疆的河阗改为和田，江西的雩都改为于都。但是，在方言地区，尤其是在东南沿海地区，还存在着大量的生僻字、方言字和异体字。对这些字如何取舍，尚需进一步研究，制定可行的规范。

第三，实行规范的地名拼音拼写法。地名在不同的语言中有不同的书写形式，文字形式不同，使用不同文字的人就看不懂。如果使用国际上通用的罗马字母来拼写地名，那就会便于国际交流。但是，我们必须首先制定自己的法定的罗马字母拼写法，否则一个国家一种拼写，也只会造成拼写上的混乱。1975 年，在联合国经社理事会在纽约召开的国际地名标准化会议上，我国代表指出"用汉语拼音字母拼写中国地名是我国的既定方针"。1977 年，在雅典举行的联合国第三届地名标准化会议上，我国代表团再一次阐明我国的主张：国际地名标准化要以各国地名的标准化为基础，各国主权范围内的地名的罗马字母标准拼写形式，应当作为国际的标准形式。在这次会议上，我国提出的采用汉语拼音作为中国地名罗马字母的国际标准提案获得通过。1984 年，由中国地名委员会、中国文字改革委员会和国家测绘局公布了《中国地名汉语拼音字母拼写规则（汉语地名部分）》，提出了较为成熟的拼写规则。1988 年，国家教育委员会和国家语言文字工作委员会公布了更加全面的《汉语拼音正词法基本规则》，至此，汉语地名的罗马字母拼写有了明确的国际标准。中国地名汉语拼音字母拼写规则主要内容为：

汉语地名中的专名和通名分写，每一分写部分的第一个字母大写，例如：

Běijīng Shì
Tài Shān
Táiwān Hǎixiá

专名和通名的附加成分，单音节的与其相关部分连写。例如：

Xīliáo Hé

Jǐngshān Hòujiē

Cháoyángménnèi Nānxiǎojiē

自然村镇名称和其他不需区分的专名和通名的地名，各音节连写。例如：

Wāngcūn

Zhōukǒudiàn

Jiǔxiānqiáo

非汉语地名本着“名从主人”的原则，按照罗马字母原文书写；非罗马字母的文字的地名，按照该文字的罗马字母转写法拼写。例如：

London

Washington

Lhasa

Hohhot

第四，编纂地名图书，巩固地名标准化成果。经过建国后尤其是改革开放后，全国的地名普查、地名标准化工作全面展开，取得了巨大的成就，全国大体上实现了地名的标准化。地名图书的编纂，是当代中国地名学的结晶，是标准地名社会化的渠道，承担着向社会推行标准地名的任务。地名图书主要有地名录、地名图、地名词典和地名志几类。

地名录是把一个省、一个地区或一个县范围内的各类标准地名编辑成册的法定地名工具书，它对反映地名规范的书写形式和读音具有重要意义。

地名图是以反映各类标准地名的地图。地名图用符号显示地表现象，用文字注明名称，每一条地名指称的地理位置形象直观，是传播标准地名最为有力的工具。

地名词典系统地介绍重要地名的标准名称、地理方位、名称来由以及有关的自然、经济、社会、文化状况。地名词典是权威、完备、准确的地名文献，对实现地名标准化和地名管理工作都有重要意义。

地名志是综合反映一个政区内地名现状和演变历史、重要标准地名自然和社会情况的较为详细的地名专志。地名志资料十分丰富详尽，对专业工作者和有关人员研究查阅都十分便利。

第五，应用地名学研究成果，为新地名科学命名。地名标准化工作不仅

要整理、规范已有的地名，还要指导新地命名和管理工作，使新地名直接步入标准化轨道。随着我国社会经济的发展，城市化的日益推进，可以预见，新地名将会大量出现，区域地名的分布会日益密集，如果对新地命名不及时加以规范指导，势必会造成新的地名问题。事实上，这一问题已经有所显现。明明是居住小区，却取名为某某花园、某某新村，名不副实；有些企事业单位因为占据了独特的或较大的区域，具有明显的指位性，因此我们常常借用为地名，比如南京师范大学、江苏省委。在开发的新区中，人们对某一单位并不熟悉，这些单位就不宜用来充当地名，但是现在这样的地名常常出现，往往使人难以准确定位。所以，在新开发的区域中，必须加强命名的管理，引入数字地名，采用序列化的道路街巷名称，这样地名就会有一定的规律，易找易记。不过，在给新地命名时，也要充分考虑到已经被民间接受的地名，否则，即使采用了数字地名，如果不被民间接纳，地名也就难以真正发挥作用。

## 第三节　汉语人名学

### 一、汉语人名的语言学特征

语言文字是人名的物质材料，人名从产生的那天起，就和语言文字结下了不解之缘。许慎《说文解字》云："名，自命也，从口从夕。夕者冥也，冥不相见，故以口自名。"意思是说，如果人们在夜间相遇，看不清楚，就需要"以口自名"，说出一个作为本人标记的符号，所以名从口从夕，是一个会意字。许慎的这段话，不仅说明了名的起源，还告诉人们名最初是以口相传的，点明了原始人名的口语性。文字是记录语言的书写符号系统，文字产生后，人名相应的也可以书写，于是人名又有了书面形式。人名的书面形式记载和传播着人类的文化，常常被人类学家称为人类文化的活化石。可以说，任何民族、任何时代、任何地域的人名都是以语言文字为物质材料的，人名首先是一种语言文字现象。但是，人名又不同于一般的语言现象，它是专名的一种，有着自身的语言学特点。

#### 1. 人名的意义

只要是人，就一定有名称，人名的核心功能就在于指称一个个特定的人。"名者，实之宾也"，人名首先是一个人特定的符号标记，这一点是认识

人名意义的基础。同时,人名也是一种社会文化现象,人名又蕴藏着丰富的社会文化内涵,不仅仅是一种简单的符号标记。所以,只有把人名的符号性和文化性结合起来,才能完整地把握人名的意义。我们把人名的意义分为指称意义、符号意义、美名意义和泛化意义 4 种类型,下面分别讨论。

(1) 指称意义

人名的指称意义就是指人名的特指性。例如,曹雪芹就是指称《红楼梦》的作者,曹操就是指称东汉末期的政治家、军事家、文学家。人名和所指称的对象不完全是一一对应的,一个名称可以指称若干人,这属于同名现象;一个人也可以有不同的名称,这属于多名现象。需要注意的是,无论是同名还是多名现象,都没有改变人名的指称特性,同名只不过是一个名称指称不同的人,在实际的生活当中,同名的绝对数量虽然较大,但在一定的交际环境当中,同名几率相对较小,一般不会引起大的误会。一人多名的情况比较复杂,比如一个人有小名、学名,小名指在小时候使用,若非亲朋好友,长大后别人大多不知道,这种情况就不存在误会的问题。再如一个人可能会有绰号,宋江绰号及时雨,李逵绰号黑旋风,吴用绰号智多星,这些名称不但不会混淆名实,反而加深了人们对宋江、李逵和吴用的认识,使人过目不忘。但有时也会引起混淆,比如鲁迅有许多笔名,有的笔名如张承禄、赵士仪、安子掌等等,使用的目的就是为了混淆视线,以避国民党书报检察官的耳目。

人名所指称的对象,不仅可以是现实或历史上实际存在的人物,还可以指称人们想像创造出来的虚幻人物。《西游记》创造了神话世界,也创造了孙悟空、猪八戒、沙和尚等等指称虚幻人物的人名世界。《红楼梦》中的贾宝玉、林黛玉、薛宝钗、王熙凤、刘姥姥等不朽的艺术形象,也使他们的名字获得了长久的生命力。

(2) 符号意义

人名的基本功能是指称,指称意义是人名的核心意义。与此相适应,人名所使用的语言材料的固有意义,在人名中逐渐虚化,成了单纯的指别符号。比如“高”本来和“低”相对,是形容词,指从底部到顶部的距离大,但是作为姓氏,“高”并没有“距离大”的意思。“熊”是“熊科动物的统称”,但用作姓氏时,“熊”姓和“熊”的词义没有必然的关系。汉语中的“第二”、“第五”、“第八”表示序数,同时也是复姓,用作复姓时和表示序数并没有必然的关系。姓是如此,名也一样。所以无意义的助词可以入名。春秋战国多有在姓名中间加一个“不”字的现象,比如韩有申不害,秦有吕不韦,晋有韩不信,

楚有任不齐，鲁有宓不齐。王引之《经义述闻》卷二十二春秋名字解诂云："鲁宓不齐，不，语词，不齐，齐也。""楚任不齐，不，语词，不齐，齐也。"还有的在姓名之间用一个"之"字，比如《左传》中有烛之武、介之推；到魏晋南北朝时，"之"的使用更是盛极一时，例如大书法家王羲之一家三代都用"之"，王羲之子辈有玄之、凝之、徽之、操之、献之，孙辈有祯之、静之。虚词如此，实词也有类似情况。宋代名字中喜欢用"老"、"叟"、"翁"，比如"胡唐老"、"王国老"、"孟唐老"、"王廷老"、"刘羲叟"、"徐清叟"、"刘辰翁"、"王积翁"，等等。所以，名字所用的字词，不宜都从理性意义上去理解，否则就会闹笑话。有一个老师最好午睡，学生问他"宰予昼寝"是什么意思，他说："宰就是杀，予就是我，昼就是中午，寝就是睡觉。宰予昼寝的意思就是杀了我也要午睡。"

(3) 美名意义

人名的指称意义和符号意义体现了人名的专名性质，是人名基本功能的体现。但是，人名还具有社会文化属性。人们不仅需要一个区别于他人的符号，而且需要一个美好的符号，因此，人们取名时总是倾向于选择美好的字眼，总是倾向于追求美好的名字。有的名字以姓连名，突破了姓名的符号意义，使姓名产生了一种特有的涵义。比如"碧野"，姓"碧"名"野"，姓名连义，取义"碧绿的原野"。"马识途"取义"老马识途"，"周而复"取义"周而复始"，"路遥"取义"路遥知马力"，类似的还有何其芳、柳青、江帆、苗得雨、梅兰芳、白凤鸣等等，不胜枚举。有的名字用字隽永，或者取自古言，比如"王利器"取自"工欲善其事，必先利其器"，"王朝闻"取自"朝闻道，夕死可也"；或者出自诗词，比如"冰心"出自王昌龄《芙蓉楼送辛渐》："洛阳亲友如相问，一片冰心在玉壶"，"张恨水"出自李煜《相见欢》："自是人生长恨水长东"。

(4) 泛化意义

一般情况下，人名只能指称个体，而不能指称个体具有的内涵特征。但是，有时候因为个体所具有鲜明的特征或者巨大的影响，从而使人名也具备了所指称人物的内涵特征，发生了专有名词向普通名词转化的现象。比如人们常常说调皮好动的孩子是"孙猴子"，这个"孙猴子"并不是指称"孙悟空"，而是指像孙悟空一样调皮好动的一类人；雷锋是一个乐于助人、品质高尚的解放军战士，1963 年毛泽东发出了"向雷锋同志学习"的号召后，因为雷锋的巨大影响，他的名字已经成了热心助人的代称。我们常常说"活雷

锋”,其中的“雷锋”并不是指战士雷锋,而是指具有雷锋一样品质的一类人。

总之,人名是专有名词,指称意义和符号意义是它的核心意义,美名意义和泛化意义是它的文化意义,前者反映了人名的语言特征,后者体现了人名的社会文化特征。

2. 人名的读音

(1) 古人姓名的读音

在阅读古籍的时候,常常会碰到古人的名字。古人名字的读音,大多数可以依照现在的读音来读,需要注意的有两种情况:

第一,古名文字可以假借,但是读音是固定的。人名最初是用来呼喊的,人名的书写形式是文字产生之后的事,人首先有固定的声音符号,然后才有固定的书写符号。先秦的时候,名字的字形和读音还没有完全稳定的搭配,读音固定,但文字可以假借,比如伏羲也写作庖羲、包牺、宓羲,皋陶也写作咎繇,伯益也写作伯翳、后益、化益,孟轲字子舆,还写作子车、子居,荀况人称荀卿,也作孙卿。直到秦汉以后,人名的音形义才逐渐定型。

第二,古人的姓名往往有约定俗成的读音,不能想当然地按照习惯的读音去读,下面是一些常见人名的读音:

皋陶(yáo)、傅说(yuè)、伍员(yún)、赵衰(cuī)、乐(yuè)毅、句(gōu)践、墨翟(dí)、范蠡(lǐ)、李适(kuò)、冒顿(mòdú)、李悝(guī)、景差(jié)、米芾(fú)、樊於期(wūjī)、共(gōng)叔段、郦食其(yìjī)、曹大家(tàigū)、高渐(jiān)离。

(2) 姓氏的读音

大多数姓氏的读音和这个字的非姓氏用法读音一致,如樊篱的樊和樊哙的樊都读 fán,树林的林和林黛玉的林都读 lín。但是,也有许多姓氏有自己独特的读音,又可以细分为以下几种情况:

第一,姓氏专用字,只有一种读音。这些字一般比较生僻,例如:

厍(Shè)、乜(Niè)、逯(Lù)、迮(Zé)、芈(Mǐ)、郝(Hǎo)、邰(Tái)、鄢(Yān)。

第二,姓氏读音是某个多音字的一个读音,例如:葛字有两个读音,当“纠葛”讲时读作 gé,当作姓氏时读作 Gě;盖字也有两个读音,分别读作 gài 和 gě,作姓氏用时,读为 Gě。类似的如纪字作姓氏用时读为 Jǐ,不读 jì;缪字作姓氏用时读为 Miào,不读 miù(纰缪)和 móu(绸缪),仇字作姓氏用时读作 Qiú,不读 chóu。

第三，某个姓氏用字是多音字，不同的读音表示不同的姓氏，例如，乐字是多音字，分别读作 Yuè 和 Lè，是不同的姓氏。类似的还有相(Xiāng)和相(Xiàng)，柏(Bó)和柏(Bǎi)。

第四，姓氏用字因为字形简化而容易混淆，例如，"郁"是"鬱"的简化字，但"郁"和"鬱"是不同的姓氏，当作姓氏用时，"鬱"不简化为"郁"。再如"于"是"於"的简化字，但是"于"和"於"是不同的姓氏，当作姓氏用时，"於"不简化为"于"。

**3. 人名的用字**

(1) 利用字形分合形成的名字

汉字大多是合体字，只有少量的独体字，所以大多数汉字可分可合，灯谜、对联常常利用汉字的这一特点，给人们带来了无限的趣味与美感。比如下面的灯谜：

目字加两点，莫当贝字猜；
贝字欠两点，莫当目字猜。

这个灯谜要求打二字，谜面把"贺(賀)"字拆为"目"、"加"和两点，把"资(資)"字拆为"欠"、"贝"和两点，谜底就是"贺"和"资"。再如下面的析字联：

品泉茶，三口白水；
竺仙庵，二个山人。

上联"品"字拆为"三口"，"泉"字拆为"白水"，下联"竺"字拆为"二个"，"仙"字拆为"山人"，可谓妙趣横生。

人们在取名时也常常利用汉字的这一特点，使得名字新颖别致，引人注目。归纳而言，有以下几种格式：

第一，姓和名使用同一个偏旁，例如江淹、流沙河、沙汀，姓、名都用了三点水作偏旁，艾芜、芦荻都是草字头。

第二，因姓构名。有的是名取姓的一部分，比如：陈东、盛成、聂耳、张弓，或者名为姓的增文，例如牛犇、石磊、林森。

第三，名为姓的分文。例如：老舍原名舒舍予，舍予即为舒的分文。张长弓、李木子、杨木易等亦然。

第四，名为姓的增益，例如魏巍、方芳、韦苇等。

(2) 人名用字有地域特点

中国姓氏分布有明显的差异，张、王、李、刘多见于北方，陈、赵、林、吴多见

于南方。陕西的党、路、权等姓，四川的卿、阳、补等姓，上海的火姓，都是该地特有的地方姓氏。各个地方的语言风俗不同，不同的地方的取名用字往往都富有地方特点。比如绍兴，人名喜欢用“根”、用“阿”，如泉根、寿根、阿伟、阿强；秦晋两地孩子的小名多用粗俗丑陋的字眼，比如铁胆、狗子、柱子。

(3)人名用字的时代性

人名的命取往往受到时代的影响，相应的，取名用字常常表现出强烈的时代风貌。萧遥天研究指出，魏晋六朝的名字中多用“之”字，五代人名喜欢用“彦”字，宋人又喜爱用表示老态的字，比如“老”、“叟”、“翁”等等。不仅古代如此，现代也是一样。在老一辈的人名当中，有很多祈求福寿，期盼荣华富贵的字眼，比如李得贵、张来福、王昌禄、赵寿丰等等。解放以后，人们的观念发生了巨大的变化，这类名字就很少了。不过，解放后历次重大的政治运动社会变革也清晰地体现在人名中。建国前后出生的孩子有很多取名为建国、解放，抗美援朝时又出现了卫国、援朝、抗美等名字，20 世纪 50 年代中期，社会主义建设高潮中又出现了建设、建强等类似的名字，大跃进的时候跃进、胜天、卫星、超英等豪迈的名字大量出现，三年困难时期，出现了坚强、自力、更生等表现人们自强不息精神的名字，“文化大革命”时期，更是出现了一批带有强烈的时代烙印的名字，比如造反、忠红、红卫、卫东、文革、永红等等，改革开放后，又出现了仿效外国人名的名字，比如纪因斯坦、沈芳娟子、殷乐笑子等。现在是网络时代，网民的名字更是千奇百怪，求新立异，诸如杀手帮主、毁灭天王、风花雪月、蜘蛛侠、疯狂老鼠等等，五花八门的名字几乎成了网络世界的一道风景。

(4) 因避讳而引起的字形变异和缺省

所谓避讳就是在言语交际中避开忌讳的字眼。避讳制度开始于周朝，是尊祖敬宗的宗法制度的产物。进入封建社会后，得到了进一步的强化，不仅要避国讳、避圣讳，而且要避家讳。避讳会使事物的名称改变，比如二十四节气的“惊蛰”原叫“启蛰”，因为避汉景帝刘启讳，改“启”为“惊”，再如为避唐太宗李世民的讳，改民部尚书为户部尚书。同时，避讳还会引起用字的变异。据陈垣考证，因避讳而使字形出现缺笔始于唐朝，盛行于宋，传至明清。有时改用同义字，比如汉明帝名叫刘庄，东汉人就把“庄”字改为“严”字，庄子也就变成了“严子”，这是因避讳而改姓的例子。有时避讳又用空字法，将应避讳的字空而不写，比如唐人在编撰《隋书》时为了避李世民讳，把“王世充”写为“王　充”，把“徐世绩”写为“徐　绩”。如果不知道这一点，非

常容易引起误会。

## 二、汉语人名系统

人人都有名字。每个人一生下来，甚至在他还没有出生的时候，长辈就会给他起名字。就汉族而言，人们往往不仅有一个名字，而且可能有好几个名字，比如小名、学名、字、号等等。不同的名字都有不同的意义和用途，它们构成了汉语复杂的人名系统。一个人不仅有名，还有姓。姓代表着一个人的家族的一种标记，它反映了中华民族重宗族、重血缘的文化传统。“姓”和“名”是中华民族人名系统的两根支柱，“姓＋名”是中华民族姓名系统最为基本的结构方式。下面从“姓”和“名”两个方面来说明汉语人名系统。

### 1. 姓氏

在现代社会，姓和氏没有什么区别，姓就是氏，氏就是姓。但在上古时代，姓和氏是两回事，含义不同，作用也不一样。那么，姓、氏的本来含义是什么？它们之间有什么关系？现代的姓氏是怎么形成的呢？

(1) 什么是姓氏

传说中远古帝王的姓都是女字旁，比如炎帝姜姓，黄帝姬姓，虞舜姚姓，这是为什么呢？古姓大致有两个来历，一是来源于图腾。比如后稷的母亲姜源因踩了熊迹而后生后稷，实际就是图腾化姓。二是根据女姓始祖生育时的地点而得姓。虞舜因为居于姚墟而姓姚，炎帝因为居于姜水而姓姜。古姓多从女字旁，说明了姓的产生可以上溯到母系氏族社会，反映了“知母不知父”的事实。在母系氏族社会中，人类按照母系血缘分成不同的氏族，每个氏族都以图腾或居住地作为互相区别的族号，这个族号就是“姓”。《说文解字》：“姓，人所生也。”这也说明姓的本义就是指源于同一母系的氏族族号。

氏的产生晚于姓，先有姓，后有氏。随着属于同一氏族的子孙的繁衍，同一母族往往分出若干支族，每个支族都有区别于其他支族的名称，这个名称就是“氏”。所以《通鉴外纪》说：“姓者，统其祖考之所自出；氏者，别其子孙之所自分也。”

进入夏商周三代后，姓氏逐渐带上了浓厚的政治和宗法色彩，《左传·隐公八年》：“天子建德，因生以赐姓，胙之土而命之氏。”意思是说，周天子分封有德的人为诸侯，根据他们始祖出生的情况给他们赐姓，赏赐给他们土地并因此而确定他们的氏。就是说，这时的姓氏都是天子赏赐的，是非同一般的，姓氏已经成了贵族的标记。与此相适应，姓氏的功能也发生了变化。之

前姓用来区别血缘，氏用来区别子孙，到了这时，男子称氏用来区别贵贱，女子称姓用来区别婚姻。

到了秦汉时代，维系周王室宗法制度的姓、氏制度土崩瓦解，姓、氏合二为一。顾炎武《日知录》中说："自战国以下之人，以氏为姓，而五帝以来之姓亡矣。"又说："姓氏之称，自太史公混而为一。《本纪》于秦始皇则曰姓赵氏，于汉高祖则曰姓刘氏。"从此之后，姓就是氏，氏就是姓，不仅贵族有姓，普通老百姓也有了姓。

(2) 姓氏的来历及种类

我们现代的姓，绝大多数可以上溯到几千年前，可谓源远流长。姓氏的来历，大体可以归纳为以下几种情形：

第一，来源于远古母系氏族的族号。姓本来是族属标记，其后人有的便直接承袭为氏，这类姓氏多为女字旁，主要有：姚、妫、姜、姬、嬴、子、任、姒、风等。

第二，以国名为氏。以此命氏的都是受封的诸侯国国君及其后裔。以周王室同姓封国的国名为氏的有鲁、晋、曹、郑、滕、虢、管、卫、吴等48氏，以周王室异姓封国的国名为氏的有齐、楚、宋、陈、赵、杞、邓、薛、徐、曾等60氏。

第三，以邑名为氏。邑是各诸侯国分给卿、大夫的封地，卿、大夫就以邑名为氏，例如刘、白、崔、卢、鲍、费、范等。

第四，以乡亭之名为氏。周代爵位分为公、侯、伯、子、男5种，相应的封地为国、邑、关、乡、亭5等，以乡亭为氏的有裴、陆、庞、郝、糜、采、欧阳等。

第五，以居住地为氏，例如傅、桥、东郭、南宫、西间、北丘、百里等。

第六，以先人的字或名为氏。诸侯国国君的儿子称公子，孙子称公孙，玄孙不能再称公孙，而以祖父的字为氏。比如宋桓公生公子目夷，字子鱼，目夷生公孙友，公孙友生鱼莒、鱼石，鱼莒、鱼石以祖父的字为氏。如果祖父没有字，就以名为氏，例如鲁孝公生公子展，公子展无字，公子展的孙子展无骇、展禽就以展为氏。类似的还有林、方、游、孔、颜等。

第七，以官职为氏，例如：司马、司徒、史、帅、钱、校、军、库、李等。

第八，以次第为氏，例如：伯、仲、叔、季、丁、癸、第二、第八等。

第九，以技艺为氏，例如：巫、屠、卜、陶、甄、匠、优等。

第十，以谥为氏，例如：敬、庄、康、武、文、幽、昭、厉等。

第十一，以事为氏，例如汉武帝的丞相田千秋因年老体弱，武帝特许他

乘车出入朝廷，时称车丞相，后人就以车为氏，类似的还有痛、蒲、白马等。

(3) 姓氏的现状

汉族姓氏的数量，经历了由少到多的发展过程。当代到底有多少姓氏？姓氏人口和地域分布有什么特点，这个问题目前还没有定说。中国科学院遗传研究所杜若甫和袁义达根据国家统计局提供的 1982 年全国人口千分之零点五随机抽样资料及 1970 年台湾省出版的《台湾地区人口之姓氏分布》一书进行统计调查，研究表明，全国汉族姓氏中李姓最多，其次为王姓和张姓。占汉族人口 1%以上的大姓共有 19 个，除了李、王、张外还有刘、陈、杨、赵、黄、周、吴、徐、孙、朱、高、林、何、郭、马，这 19 个大姓占汉族人口的 55.60%。

汉族姓氏虽然数量众多，几乎日常所用的 3 000 多个汉字都是姓氏，但是，研究表明，最常见的 100 个姓氏总计约占汉族人口的 87%，它们是：

李、王、张、刘、陈、杨、赵、黄、周、吴、徐、孙、朱、高、林、何、郭、马、罗、梁、宋、郑、谢、韩、唐、冯、于、董、萧、程、曹、袁、邓、许、傅、沈、曾、彭、吕、苏、卢、蒋、蔡、贾、丁、魏、薛、叶、阎、余、潘、杜、戴、夏、钟、汪、田、任、姜、范、方、石、姚、谭、廖、邹、熊、金、陆、郝、孔、白、崔、毛、邱、秦、江、史、顾、侯、邵、孟、龙、万、段、雷、钱、汤、尹、黎、易、常、武、乔、贺、赖、龚、文

**2. 名字**

古人幼有名，长有字，死有谥，姓名制度远比现在复杂，一个人有若干个称号是很正常的。比如杜甫姓杜名甫，字子美，因为曾经在长安附近的杜陵寄居，所以又称杜陵、杜少陵。因为久未入仕，所以又自号少陵野老、杜陵布衣。安史之乱时曾任左拾遗，又称杜拾遗。入川后任检校工部员外郎，又称杜工部。晚唐还有一位诗人杜牧，为了表示区别，世人又称杜甫为老杜。那么，我们的姓名制度是怎样的？都有哪些种类？名、字、号之间又有什么关系呢？

(1) 小名

小名，也叫乳名、奶名。小名古已有之，曹操小名阿瞒，刘禅小名阿斗，司马相如小名犬子，陶渊明小名溪狗。小名是昵称，洋溢着父母的宠爱，因为不是正式的名字，所以取名随意，有的以产次命名，如阿四、小三子，有的以动物命名，如小狗、小熊、大猫、虎子、牛牛等，有的以花卉命名，如小花、蕾蕾，类别难以尽举。小名的另一个特点取名粗鄙，比如希望孩子长得俊俏，却反其意而取名为丑蛋、丑丑，希望孩子聪明伶俐，却取名为傻傻、憨子，希

望孩子强壮健康，就选用结实的东西给孩子命名，比如石头、铁蛋、铁柱、钢钢、金锁等。

(2) 大名与字

小名不登大雅之堂，长大后大多要更换一个正式的名字，这个名字就是大名。大名和小名有明显的区别，小名是昵称，所以随便、亲切，大名是一个人正式的名字，所以比较庄重、严肃，具有丰富的含义。大名的命取丰富多彩，难以尽举。一般来说，大名大多是由父母长辈命取的，有的寄托着父母长辈的希望，比如自立、有为、昌乐、昌来、维贤、慎行、崇理、理嘉、福义、世英、祥明、瑞敏、瑞民、建华、志伟、克勤、达仁、贤纯、英哲、行健等，这一类占绝大多数。有的是纪念某一时刻、某个地方，比如迎年是纪念出生的时间，表示在过年前出生的；《红楼梦》里的元春就是在大年初一出生的，所以叫元春；而春华、夏雨、晓冬等名字纪念的是孩子出生的节令。鲁迅的儿子叫周海婴，是因为他出生在上海。著名歌唱家关牧村出生在河南省新乡市牧野村，所以取名为牧村。还有的是以父母的原籍命名的，比如文学评论家秦晋，父亲是山西人，母亲是陕西人，恰好又姓秦，真是恰如其分。类似的比如王沪宁、朱湘蓉、李申渝、张京海等等，从这样的名字中一般能推断出这个人父母的原籍。孩子是爱情的结晶，有的父母就用两人的姓氏给孩子取名，以作纪念，比如王武丽（父姓王，母姓武）、周徐才（父姓周，母姓徐）、杨柳（父姓杨，母姓柳）。有的父母文学修养比较高，往往从古典诗词中摘选字词，给孩子取一个隽美的名字，比如谢白羽，“白羽”取自卢纶《塞下曲》：“平明寻白羽，没在石棱中。”还有的是以抓周的习俗来给孩子命名的，比如我国著名学者钱钟书就是在抓周的时候因为抓了一本书而得到“钟书”这个名字的。

当然，有些大名是自己改动而成的，这些大名或者寄托着某种感情，或者表现了理想抱负。郭沫若原名郭开贞，他不喜欢这个名字，于是取家乡的“沫水”和“若水”各一字作为自己的名字，表达了对家乡的思念。著名的教育家陶行知原名陶文濬，后来因为信仰王守仁“知是行之始，行是知之成”的思想，改名为陶知行；后来进一步认识到真知出自实践的道理，因而又改名为陶行知。胡适原名洪骍，后来受到达尔文进化论“物竞天择，适者生存”思想的影响，改名为胡适。

古人不仅有姓有名，而且还有字。现在所说的名字，对应于古代的名，古人除了有一个正式的名之外，还要有一个字。字是由名演化而来的，因而现在统称名字。

在古代，称名用字大有讲究。《礼曲礼》："男子二十冠而字，女子许嫁笄而字。"名幼年就有了，用来分别彼此；字是成年后加上去的，以示尊重。在古代，名用于自称，表示谦虚。只有尊长对晚辈、上级对下级才可直呼其名。晚辈对长辈、下级对上级、平辈称对方尊长，都要称字，表示敬重；平辈之间也可以彼此称字，表示礼貌。

名和字的关系非常密切。《说文解字》："字，乳也。"字的本义是乳育的意思，引申为孳乳增多的意思。所谓字，就是由名孳乳而来，是名的延伸，因而《颜氏家训》云："古者名以正体，字以表德。"总括而言，名和字的关系大约有以下几种。

1）名与字意义相同或相近。屈原本名平，原是字，平和原同义；诸葛亮字孔明，亮和明同义；张辽字文远，辽和远同义，东方树字植之，树和植同义。陆游字务观，秦观字少游，游和观，意义相近。司空曙字文明，曙和明义近。

2）名与字意义相反。韩愈字退之，公孙黑字子皙，赵衰字子余，晏殊字同叔，其中愈和退、黑和皙、衰和余、殊和同，都是反义词。

3）名与字意义相关。孔丘字仲尼，丘即山，孔子生于尼山，排行第二，所以字仲尼。周瑜字公瑾，瑜、瑾都是美玉。山涛字巨源，由山中的波涛联想到巨大的源头。赵云字子龙，"云从龙，风从虎"，云、龙由此相关。周鼎字伯器，鼎为器类，由此相关。苏轼字子瞻，轼是古代车前面用作扶手的横木，瞻即视，《左传·庄公十年》："登轼而望之。"两者相关。刘过字改之，显然是有过则改的意思。

4）名、字取自古语、成语。钱谦益字受之，取自"谦受益"；高明字则诚，取自"明则诚"；薛玉壶字冰心，取自"一片冰心在玉壶"；陈群字人鹤，取自"鹤立鸡群"；潘岳字安仁，取自"仁者乐山"，以岳代山。陆羽字鸿渐，取自《周易》："鸿渐于陆，其羽可用为仪。"

5）以伯、仲、叔、季为字。晋荀林父字伯，鲁公子庆父字仲，秦范雎字叔，鲁公子友字季。伯仲叔季还常常加在一言字之前构成二言字，比如鲁孔鲤字伯鱼，孙权字仲谋，鲁孔纥字叔梁，晋张翰字季鹰，班固字孟坚。

6）拆名为字。宋玫字文玉，明刘侗字同人，清舒位字立人，其他如姚椿字春木，蒋伊字尹人，毛奇龄字大可。

7）表示仰慕。颜之推字介，字和名合起来就是介之推，表示对介之推操守的敬仰，类似的有雷渊字希颜，宋学朱字用晦，章如旦字希周，袁尊尼字鲁望。

(3) 号

古人在名、字之外的往往又自命一个别号，比如李白号青莲居士，陆游号放翁。如果说“名以正体”、“字以表德”的话，那么号就是寄托心声，是古人用来抒发自由意志的一种方式。号起源很早，春秋时已经有了称号的现象，但直到唐宋以后，号才开始流行，明清时达到鼎盛，上至帝王，下至庶民，都爱取号。

古人取号，形式自由，少的有两字，多的可达十几字，不过以两字、三字和四字最为常见。用字也很自由，不过也并非无迹可循，萧遥天把取号的用字分别归类，可见古人取号的大端。举要如下：

1) 表示称谓的有叟、老、翁、子、生、人、民、客、士、主、公、君、长、友、夫、臣、郎、史、父、徒、先生等。例如晋陶潜号五柳先生；唐陆希声号君阳遁叟，杜甫号杜陵野老，韩偓号云山樵人，郗纯号伊川渔父，张志和号烟波钓徒；宋陆九渊号象山翁，陈抟号扶摇子，杨璞号东野遗民，黄升号花庵词客，欧阳修号六一居士，李椿年号逍遥公，陈翥号桐竹君，邓林号四清社友，柳开号东郊野夫，米芾号海岳外史；明王世贞号五湖长，左光斗号浮丘生；清罗振玉号楚雨楼主，高翔号山林外臣，金农号金二十六郎。

2) 表示身份的有布衣、僧、和尚、仙、道、隐、逸等。例如唐杜甫号杜陵布衣、贾岛号浪仙，宋胡仔号苕溪渔隐，元赵孟頫号松雪道人，明陈洪绶号云门僧，明李贞号大呆和尚，清钱万里号南村诗逸。

3) 表示百业的有渔、农、耕、樵等。例如清樊曙江号东溪老渔，清吴非号古牛耕者，清潘荫东号五山寄樵，宋陈世昌号大田农。

4) 表示居处的有斋、堂、庵、阁、庄、室、屋、舍、庐、房、馆等。例如宋杨万里号诚斋，宋韩琦号昼锦堂，宋朱熹号晦庵，宋顾禧号漫庄，唐刘禹锡号陋室，明王冕号梅花屋，明陈道复号五湖田舍，清王夫子号败叶庐，明胡应麟号少室山房，清王文治号丁香馆。

5) 表示山川河海的有山、峰、谷、溪、池等。例如宋陈师道号后山，宋黄庭坚号山谷，宋史达祖号梅溪，明徐渭号天池，元黄公望号乙峰。

6) 表示园林的有园、林、墅、榭等。例如元虞集号道园，胡梦昱号竹林；清胡湄号松风别墅，黄安涛号枕山榭。

7) 表示器物的有轩、窗、船、舟、衣、瓢等。例如宋辛弃疾号稼轩，吴文英号梦窗，程垓号书舟；元殷奎号春水船；清朱卉号草衣，王夫之号一瓢。

号所寄托的含义是丰富多彩的，有的是志记癖好，如明朱国盛号拜石

斋，清顾谦号爱梅轩，清金农号二百砚田富农；有的以号明志，如顾炎武号避青先生，以示抗清之志；贺知章号四明狂客，陆游号放翁，邵雍号安乐先生，则表明了独特的性情与心态；有的描述形貌，比如明祝允明号枝指生，清金霖号左手道人，邓大林号长眉道人；有的表明环境，比如陶渊明号五柳先生，欧阳修号六一居士；有的抒发性情，如唐司空图号知非子，宋刘晋号求诸已斋；有的自勉，比如杨万里号诚斋，勉励自己为人为学要正心诚意。

(4) 地望名、宦地名和官爵名

地望名是用出生地称人的一种名号，比如李鸿章出生于安徽合肥，所以称李合肥；王安石出生于江西临川，世人称之为王临川。地望名大致始于唐宋，首先在知识分子中使用，一般都是社会上的名流。地望名的格式为"姓＋地名"，略举数例：唐柳宗元是山西河东(今山西永济)人，人称柳河东，他的作品集也叫做《柳河东集》；清张之洞是直隶南皮人，世称张南皮；康有为是广东南海(今广州)人，人称康南海，作品集也称为《康南海先生诗集》。

地望名中的地名一般是指一个人出生的地方，但是，在封建社会，人们有浓厚的郡望意识(所谓郡望意思就是世居某郡为人们所仰望，指魏晋至隋唐时每郡的显贵家族)，推崇高门大姓，注重门第是普遍的心理。所以，地望名有时就不是一个人的出生地，指的是郡望。比如唐韩愈的出生地在河阳(今河南孟县)，因为当时韩姓以昌黎为郡望，所以韩愈自称昌黎人，他的传世之作也叫做《韩昌黎集》。

宦地名是以某人的任职地相称的一种名号，格式也是"姓＋地名"。宦地名中的地名，一般来说，是对一个人来说比较重要的地点。比如西汉贾谊才能出众，却被贬为长沙王太傅，世人称之为贾长沙；东晋的陶潜虽然担任彭泽(今江西彭泽)令只有 80 多天，但这正是他喜爱自然、性喜自由的体现，所以世人称他为陶彭泽。唐柳宗元长期被贬柳州，在柳州大兴善政，造福一方，世人称为柳柳州。

官爵名是指以某人的官衔、爵位称呼的一种名号，格式是"姓＋官衔"或"姓＋爵位"。比如杜甫曾经担任检校工部员外郎，世称杜工部；嵇康官拜中散大夫，世称嵇中散；王羲之官至右军将军，世称王右军。汉班超出征西域，功盖当时，被封为定远侯，人称班定远；魏徵封郑国公，世人称为魏郑公；王安石被封荆国公，人称王荆公。

(5) 笔名

所谓笔名，就是发表文章时所署的别名。笔名大约起源于元末明初。

从明代开始，话本小说等通俗文学样式开始兴起，这些作品被看作是不登大雅之堂的雕虫小技，所以文人们在署名时就不愿意签署真实的姓名，而是另外取一个化名来刊行。比如《彭公案》题名为“贪梦道人”，《海上花列传》题名为“花也怜侬”。当然，也有的是为了掩人耳目用笔名，比如《金瓶梅》的作者题名为“兰陵笑笑生”，或者是因为躲避文祸而签署别名，比如晚清谴责小说《老残游记》题名为“洪都百炼生”。“五四运动”后，使用笔名蔚然成风，取代了古代的字号别名，发展为现代知识分子主要是作家的习用别名。

现代作家的笔名形形色色、丰富多彩。比如鲁迅，据考证一共用过 140 多个笔名，形式多样，含义深隽。有的寄托乡情，比如越客、越丁、越侨，说明自己是越人（绍兴春秋时属越国）；有的用于自励，比如孺牛即“俯首甘为孺子牛”的意思，鲁迅即迅捷力行的意思；有的用于自谦，比如“巴人”即下里巴人的意思，“不堂”是不登大雅之堂的意思；有的讽刺论敌，比如“尊古”意在讥讽尊孔复古的守旧派，“封余”意在反击论敌攻击鲁迅是“封建余孽”；还有的取自典故，寄托深义，比如“庚辰”之名出自唐李公佐撰《古岳渎经》，庚辰是神话中的英雄，鲁迅以此为笔名，表达了不畏强暴、为民制邪的理想。

（6）谥号

谥号是封建时代在帝王公卿诸侯死后根据其生前的事迹而给予的评定褒贬的一种称号。谥号起源于周代。上古有号无谥，到了周朝，生有名，死有谥，帝王死了之后，一定要避讳他的名号，所以谥号产生之初，主要是用来避讳的。因为是死后加的名，所以谥号又是盖棺定论，总结一个人生前的行为，用一两个字评定褒贬。比如姬昌谥为文王，姬发谥为武王。“经纬天地曰文，慈惠爱民曰文”，“刚强里直曰武，威疆睿德曰武”，以文、武颂扬他们的功绩；姬胡谥为厉王，姬宫谥为幽王，“杀戮无辜曰厉”，“动祭乱常为幽”，以厉、幽贬斥他们的作为。

除了秦朝短暂地废除了谥法外，从周代产生到清代退出历史舞台，谥法一直延续了 3 000 多年，谥法的主要内容有：

第一，评定一个人的行为，比如表彰有德的有恭（尊贤敬让曰恭）、定（纯行不二曰定）、德（谏争不威曰德）、襄（甲胄有劳曰襄）、成（安民立政曰成）等。贬斥失德的有缪（名与实爽曰缪）、戾（不悔前过曰戾）、丑（怙威肆行曰丑）、夸（华言无实曰夸）、炀（好内怠政曰炀）等。

第二，评定一个人的际遇，比如际遇好的有夷（安心好静曰夷）、静（宽乐

令终曰静)、康(丰年号乐曰康)等;际遇坏的有怀(慈仁短折曰怀)、愍(在国逢难曰愍)、悼(恐惧从处曰悼)等。

上面讲的都是公谥,除了公谥之外,还有私谥。所谓私谥就是不经过朝廷的议定而由乡里亲属或门人好友所定的谥号。私谥起源很早,春秋时展禽死后由妻子定谥为惠(爱民好与曰惠,柔质慈民曰惠),因为居于柳下,所以展禽被称为柳下惠。再比如陶渊明去世后由好友颜延之定谥为"靖节征士"。

3. 名字现状

周秦以来,汉族人名逐渐形成了一个包括名、小名、字、号、地望名、官爵名、谥号、笔名等在内的复杂系统,可以说,汉族人名系统是漫长的封建文化的产物,蕴含着丰富的历史文化内涵。在当代社会,这个系统存在的社会、文化因素已经发生了根本的改变,因而,这个系统也发生了很大的变化。有的消亡或者萎缩了,比如地望名、官爵名、帝王贵族的谥号随着封建社会的结束已经消亡;字、号在五四前后还十分常见,但是在当代,已经很少有字和号了。随着这些名字的消失和萎缩,当代人们的称呼也发生了变化。一般人之间可以直接称呼姓名,朋友之间称呼经常省去姓氏,以示亲密。与字、号不同,笔名在现当代却取得了很大的发展,笔名的使用范围进一步扩展,不仅作家诗人可以用笔名,而且有机会发表文章的人大都喜欢用一个笔名。同时,在当代大部分孩子都有小名,这大概与现在大多是独生子女,父母更加疼爱自已的孩子有关。总的来说,当代的人名系统已经大大简化,人名的符号性进一步加强,所蕴含的文化内涵有所减弱。

同时,当代人名还存在着令人担忧的重名现象。根据国家语言文字工作委员会对百万人口人名的抽样统计的结果,重名率高达25%(同名不同姓)。当代人名之所以重名率较高,首先是因为单名比例的增加。据1984年国家语言文字工作委员会对174 900人的统计,单名的比例是12.2%,有21 400人;据1989年对570 822人的统计,单名的比例是8.7%,有49 405人。单名的变化有限,重名率一定会很高,据北京市公安局统计,叫张力的有4 500多人,叫张颖的有4 000多人。其次,重名的另外一个重要原因是选用的常用字过于集中,比如女性大多使用秀、玉、兰、芳、珍、丽、淑、华、凤、梅、红、美、云、琴、霞、萍、燕、艳、娟等字,男性多用强、志、国、军、荣、伟、建、庆、明、成、忠、平、勇、刚、斌、亮、杰、辉、鹏、毅、超等字,这样,势必造成比较高的重名率。

## 三、人名的规范

当代社会是一个信息化的社会，语言文字的标准化是时代的迫切要求。人名是指称个体的专有名词，数量巨大，如果取名没有一定的规范，就会给人们的交流造成不便，就会给信息处理、户籍管理、出版印刷等方面造成很大的困难。可以说，随着现代科学技术的应用和语言文字计算机标准化处理的普及，汉语人名中的不规范现象越来越突出，给公安、邮电、人事等部门的工作和计算机信息处理带来许多麻烦，影响了工作效率。因此，人名实现规范化有重要的现实意义。目前，人名中的不规范现象主要表现为以下两个方面：

第一，人名中使用异体字、生僻字。例如有的人名不用“哲”而用“喆”，不用“村”而用“邨”，不用“杰”而用“傑”；有的为了避免重名，特意使用生僻的汉字，比如：张彧、施赟超、王熇、陈劼。这类用字妨碍交际，降低了文字信息处理的效率。

第二，名字格式随意，有的使用四字格，有的使用五字格，甚至自创复姓、仿效洋名，如“齐田美惠子”、“沈芳娟子”等，这些都违反了标准化、规范化的原则，不利于人们交流。

实现人名的规范化，要从下面几方面入手：

第一，限制人名用字，实现用字规范化。根据中国文字改革委员会(现称国家语言文字工作委员会)和国家标准局对从1977～1982年间的社会科学和自然科学的13 800万字的材料中抽样1 108万字进行的计算机统计，前3 000个高频度汉字的累计出现频率是99.15％。如果累计汉语常用的6 000多汉字的出现频率，基本上接近100％。如果我们把人名用字限制常用到6 000多字的范围之内，就可以规范取名的用字问题，这样，无论对人们之间的交流还是对信息处理都会带来很大的方便。

第二，规范取名字格，提倡使用双字格。中国人称呼姓名经常用两个字或3个字，偶尔有4个字，超过5个字就拗口，这是汉语韵律的体现。另外，在中国的姓氏中，前100个大姓占汉族总人口的87％，如果多用单名，重名率过高；如果绝大部分人起双名，再加上单名，重名率就会大大降低。因此，应该提倡起双名，尤其是大姓更要起双名。但如果是复姓，应该提倡起单名，因为汉语人名趋向于三字格，而且复姓基本都是小姓，所以起名时即使用单名，重名率也不会太高。

第三，人名不仅要有规范的形式，做到易读、易认、易写，而且还要注重命名的语言文字规律和名字所蕴含的文化内涵，努力命取美名。在日常生活中，名字主要是用来呼叫的，所以名字首先要做到声音顺口响亮悦耳。有些名字不注重声韵调的配合，读起来像绕口令一样。例如王广光这个名字，韵母都是 ang，3 个字叠韵，广和光还是双声，王是平声，广是上声，光又是平声，读起来十分拗口。因此，取名时要特别注意声韵调的配合，避免双声和叠韵，在声调的选择上，要注意和谐，平仄搭配，才能取得抑扬变化的效果。著名演员王丹凤原名王玉凤，玉和凤都是去声，读起来就很平淡，而丹是平声字，凤是去声字，两者搭配，声音就十分响亮悦耳。同时，汉语里同音字很多，所以在选择取名用字时一定要注意谐音，以免引起不雅的联想。例如朱与猪、奋与粪是同音字，如果取名朱奋，光看字的话，是个不错的名字，但是读出来就不太雅；又如有个姓吴的朋友给女儿取命“吴欣”，谐音“无心”，又有一对姓肖、姓任的夫妇，为了表示相爱情深，给孩子取名为肖任晶，谐音“小人精”，都是没有注意到谐音可能引起的不雅的联想。其次，要认真考虑用字的意义，求雅避俗。例如洁凝、敏达、友兰、潜志、壮飞等都是意义雅致的好名字。要做到这一点，除了要提高命名者文化修养外，在取名时还要注意下面几点：第一，要避免字眼的搭配意义过于肤浅，比如钱勦来、王来福、马千乘、房常有等诸如此类的名字就十分俗气；第二，要避免时代、政治色彩浓厚的字眼，比如“文革”当中卫红、保彪、跃进、反帝等名字现在看来就十分不适时宜；第三，男女是平等的，但是，这并不意味着要抹煞男女之间的性别差异，恰恰相反，在取名问题上，一定要注意性别差异，因为在某种程度上说，姓名表达了人们最基本的社会角色的区别——男性和女性。所以，男性应该选用诸如浩、超、洪、建、扬等有阳刚气息的字眼，女性应该选用玉、淑、芬、晶等有温柔气息的字眼。最后，还要注意选用汉字的形体搭配，使名字写出来美观大方。汉字大多是合体字，有的是左右结构，有的是上下结构，还有包围结构，互相搭配，往往能取得匀称活泼的效果。

## 参考书目

1. 曾世英. 中国地名拼写法研究. 测绘出版社，1981
2. 辽宁省地名委员会等编. 地名学研究(一). 辽宁人民出版社，1984
3. 中国地名委员会办公室编. 地名学文集. 测绘出版社，1985

4. 褚亚平主编. 地名学论稿. 高等教育出版社,1986
5. 王维屏. 中国地名语源. 江苏科学技术出版社,1986
6. 褚亚平等. 地名学基础教程. 中国地图出版社,1994
7. 李如龙. 汉语地名学论稿. 上海教育出版社,1998
8. 徐俊元等. 贵姓何来. 河北科学技术出版社,1985
9. 萧遥天. 中国人名的研究. 国际文化出版公司,1987
10. 王泉根. 华夏姓名面面观. 广西人民出版社,1988
11. 王建华. 文化的镜象——人名. 吉林教育出版社,1990
12. 高剑峰编著. 中国 100 个大姓. 甘肃人民出版社,1998
13. 张联芳主编. 中国人的姓名. 中国社会科学出版社,1992
14. 何晓明. 姓名与中国文化. 人民出版社,2001

# 后　记

想编一本应用语言学的教材是我们几位同仁议论多年的事了。记得还是 20 世纪末，当时肖奚强、张亚军等人还都在上海师范大学攻读博士学位，平时我们几个经常聊天，聊得最多的自然是现代汉语语法，是现代汉语语法的研究和教学。在感慨上海师范大学语法学科发展迅猛的同时，也深深地为这所学校应用语言学各分支学科发展的迟缓而叹息。当然，叹息是表象，叹息的背后是深入的思考和激烈的讨论，思考和讨论换来的是两项重大的决定，其中之一就是要编一本既适合中文系、对外汉语系、语言学系、应用语言学系本科生选修课使用，又适合语言学及应用语言学、汉语言文字学专业硕士研究生基础课使用的"应用语言学"教材。

但是，真正开始编这本教材，离当初作决定的日子已经过去很多年了。其间，肖奚强、张亚军二君离开上海，回到各自工作的学校；留在上海师大的几位又忙着各自手头的活，把这件事给耽搁了。是复旦大学出版社韩结根先生"独具慧眼"，认准了这本《应用语言学纲要》，不断地打电话，催着我们编写。韩结根先生的诚心和热心，我们在编写《新编现代汉语》时就已感受到，这次更加深了印象，没有他的催促，他的"威逼"，他的"恐吓"，《应用语言学纲要》断然不会"呱呱落地"，说不定还会"胎死腹中"，还会再一次流产。

编"应用语言学"教材实际上是非常难的。广义的应用语言学研究范围极广，几乎涵盖了语言学知识在其他科学领域的一切应用。要把所有的"应用"全部编在一本教材中显然是做不到的，我们这几个专治汉语语法的"门外汉"也没有这样的能力和功力。"写什么"的问题一直困扰着我们，在编写教材大纲的那些日子里，我们几个人互相之间打了无数次电话，每个人都为中国电讯事业的发展作出了巨大的贡献。

冯志伟先生在他的《应用语言学综论》一书中说过："应用语言学的应用因时代的不同而不同，因地区的不同而不同，因实际需要的不同而不同。"这

句话最后成了我们编写教材大纲的理论基础。首先，我们这本教材是为非外语专业的学生编的，因此，编写的内容要与汉语的应用有关；其次，我们这本教材不是面面俱到的"通论"，而是专题性的"导论"，因此，不注重体系上的完整；第三，我们这本教材要受到课时的限制，大致在36～48课时之间要把这些内容讲完，因此，教材的篇幅不能过大。在上述原则指导下，我们编就了这本教材的大纲，形成了现在这样的章节安排。

要说明的是，作为教材，这本书会在很多地方、很多问题上借用其他教材的说法，除特别需要作的说明外，一般不再注明出处，在这里我们一并致谢了。由于功力、能力、材料、时间等诸因素的影响，《应用语言学纲要》一定会有许多不足的地方，恳请专家学者和读者们向我们指出教材的缺陷，以便在今后的使用中作出修订。

这本教材最初说好不设主编，就由我们两人和肖奚强、张亚军二君编写。后来，参加编写的人多了起来，编写者一下子写上七八个似乎不太合适，肖奚强、张亚军二君又打来电话，一定要我们两人当主编，说这样才像教材的样子。我们也只好顶着"冒充"的嫌疑，当上主编。本书各章节的主要编写者为：

前言：陈昌来；第一章：陈昌来、胡建锋；第二章：肖奚强；第三章：齐沪扬；第四章：张亚军；第五章：齐沪扬、姚占龙；第六章：陈昌来；第七章：张亚军；第八章：肖奚强、张小峰。全书最后由我们两人统稿。

当这本教材出版的时候，我们两人都已到上海师范大学对外汉语学院工作了。可以这么认为，我们两人今后主要的研究方向将转为对外汉语。上海师范大学为对外汉语事业搭建了一个舞台，我们要在这个舞台上唱好歌，演好戏，为对外汉语事业作出贡献。这其实就是当初我们所作的两个决定中的另外一个。

齐沪扬　陈昌来

甲申年正月初一于上海师范大学

［Ｇｅｎｅｒａｌ　Ｉｎｆｏｒｍａｔｉｏｎ］
书名＝应用语言学纲要
作者＝
页数＝３１０
ＳＳ号＝０
出版日期＝

封面
书名
版权
前言
目录
正文